J. Differential Equations 252 (2012) 5923–5950

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Existence and differential geometric properties of continuous families of periodic three-body motions with non-uniform mass distributions

Mahdi Khajeh Salehani [a,b,*]

[a] Dept. of Mathematical Sciences, NTNU, NO-7491 Trondheim, Norway
[b] Mathematics Section, ICTP, Strada Costiera 11, I-34151 Trieste, Italy

ARTICLE INFO

Article history:
Received 3 May 2011
Revised 6 February 2012
Available online 21 March 2012

Keywords:
Newtonian three-body problem
Equivariant differential geometry
Analytic continuation method
Vanishing angular momentum periodic orbit
Non-uniform mass distribution

ABSTRACT

Using the method of analytic continuation in an equivariant differential geometric setting, we exhibit two interesting families of vanishing angular momentum periodic orbits for the Newtonian three-body problem with non-uniform mass distributions having two equal masses which connect at the celebrated figure-8 orbit, exhibited by A. Chenciner and R. Montgomery (2000) in the case of equal masses, and yield a continuous family of periodic three-body motions in the plane.

At one end of the family, when the two equal masses are infinitesimal and the third one reaches the value of $+1$, we arrive at a solution of a double Kepler problem; at the other end of the family, when the third mass is infinitesimal, we have a special case of periodic solution of a restricted three-body problem.

1. Introduction[1]

The quest for periodic orbits in celestial mechanics started long before the official date of birth of the field, which is largely agreed to be 1687, the year of publication of Newton's *Principia mathematica*. This is not surprising if we think that celestial mechanics grew as an abstract branch of mathematics inspired by the motion of celestial bodies, especially of those of the solar systems. The work of ancient astronomers, of which Ptolemy is the most important, followed by that of Copernicus and Kepler, to

* Correspondence to: Dept. of Mathematics, University of Bergen, P.O. Box 7800, N-5020 Bergen, Norway.
E-mail address: salehani.math@gmail.com.

[1] The first three paragraphs are mostly quotations from Florin N. Diacu who wrote them in the mathematical review of [3] at *www.ams.org/mathscinet/* (see [5] for the complete bibliography).

0022-0396/$ – see front matter
http://dx.doi.org/10.1016/j.jde.2012.03.005

name only the most prominent ones, shows a keen interest in describing the periodic orbits of the planets in order to predict their motion with the help of the mathematical techniques existing at that time.

At the beginning of the 18th century Newton and Johann Bernoulli were the first to prove the existence of elliptic orbits for the Kepler problem. About half a century later, Euler and Lagrange put into evidence two remarkable periodic solutions of the 3-body problem: the relative equilibria derived from collinear and equilateral central configurations [10]. It soon became clear that central configurations generate periodic solutions, so this method got established as one of the standard tools for discovering new periodic solutions.

Poincaré considered the problem of finding periodic orbits as essential not only in celestial mechanics, but for Hamiltonian systems as well: they are "the only opening through which we can try to penetrate the stronghold" [11]. He extended many of the results regarding the n-body problem to the general equations of dynamics. He also came up with qualitative methods for proving the existence of periodic solutions, without necessarily finding them explicitly; and proved that most of them are actually chaotic. From the viewpoint of Deprit and Henrard [4], periodic orbits constitute the "skeleton" around which orbits in general are organized.

Nevertheless, apart from the rather trivial family of shape invariant periodic solutions discovered by Euler and Lagrange around 1770, no new solution had been rigorously established until the figure-8 was discovered and investigated by A. Chenciner and R. Montgomery in the case of equal masses (cf. [3]).

Among many methods which can be used to find periodic orbits, such as the application of the fixed point theorem, methods of power or Fourier series, and variational methods, to name only the most popular ones, one of the most practical methods with a simple principle is the method of analytic continuation in which we need to begin our search from some already known periodic orbit.

In this article, using the method of analytic continuation (MAC) in an *equivariant differential geometric* setting (see [1,8,9] for more details about the geometric setting), we are questing for continuous families of vanishing angular momentum periodic orbits for the Newtonian three-body problem with non-uniform mass distributions in the plane. It is proved that there exist two continuous families of such periodic three-body motions with non-uniform mass distributions having $m_1 = m_2$ and keeping sum of the masses to be $+1$, which connect at the celebrated figure-8 orbit in the case of equal masses and yield a continuous family of periodic orbits.

Moving along the line segment $m_1 = m_2$ in the mass space $m_1 + m_2 + m_3 = 1$ with respect to (the heavier or lighter) m_3 deforms the figure-8 orbit in a nice and beautiful fashion which yields three different orbits: m_3 moves on an 8-shaped orbit $\boldsymbol{b}_3$, and the two equal masses move on different deformed ones which are reflectional image of each other with respect to the symmetry axis of $\boldsymbol{b}_3$ connecting its two lobes' mid-points (see Fig. 1).

Using some Computer Algebra Systems, we can start from the celebrated figure-8 orbit and modify the three masses in the mass space $m_1 + m_2 + m_3 = 1$ keeping $m_1 = m_2$ whose corresponding mass distributions with respect to (the heavier or lighter) m_3 are dual to each other. For any given mass distribution and among all possible three-body motions, MAC only guarantees the existence of the associated periodic shape curve γ^* on the shape space M^* (regarded as a 2-sphere) which meets the canonical symmetry plane (containing the Euler point e_3, the binary collision point b_{12}, and the north pole on M^*) orthogonally two times on opposite hemispheres of the 2-sphere M^*. The cornerstones of the proof of periodicity of the orbits in the inertial plane are the Hsiang–Straume's unique parameterization and monotonicity theorems, as well as their kinematic version of the Gauss–Bonnet theorem (cf. [8,9]); in the latter one, we make a beneficial use of the Lambert azimuthal area-preserving projection map from the 2-sphere M^* to an open disk in the plane.

At one end of our family of periodic orbits, when the two equal masses are infinitesimal and the third one reaches the value of $+1$, we arrive at a solution of a double Kepler problem which is a shape invariant one; and at the other end of the family, when the third mass is infinitesimal, we have a special case of periodic solution of a restricted 3-body problem which is also shape invariant. In

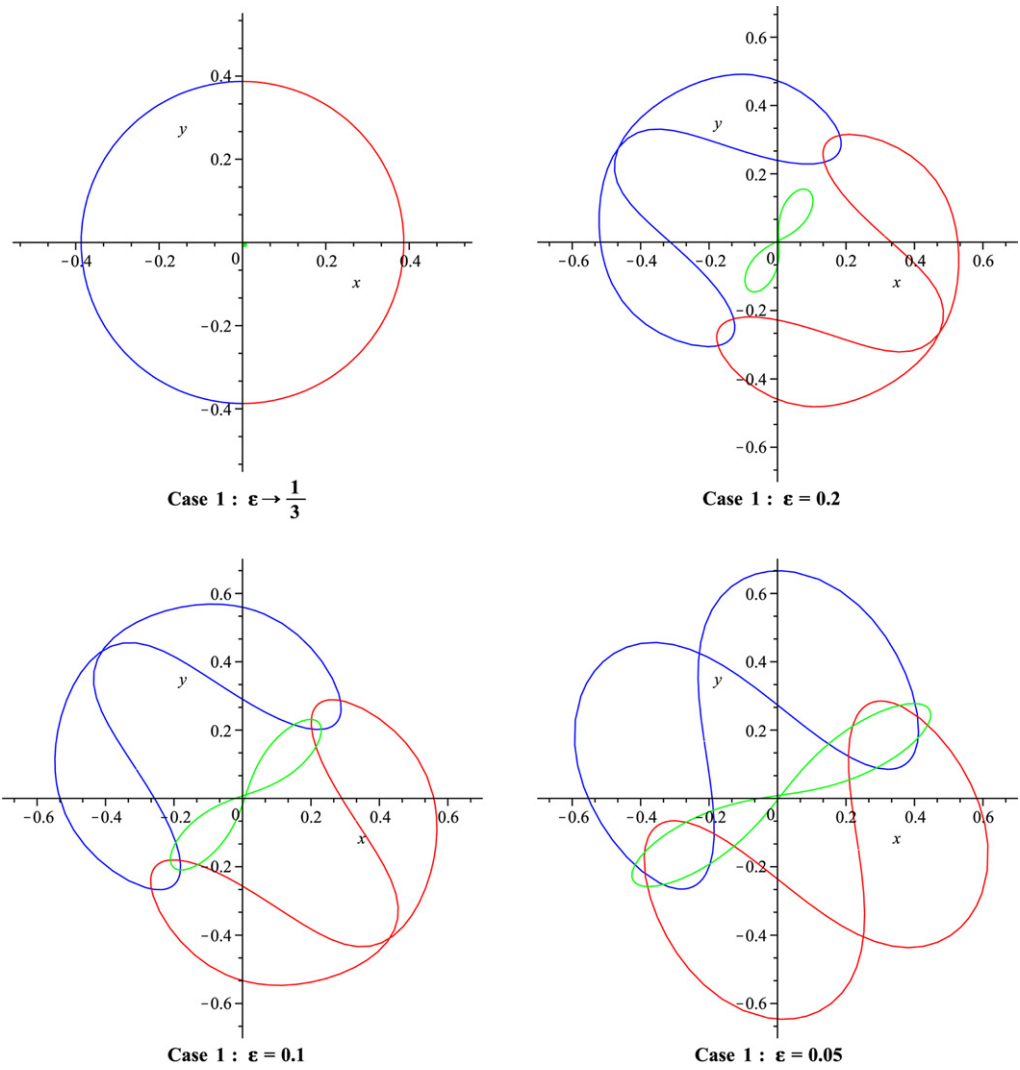

Fig. 1. Some members of our family of periodic orbits. Case 1: $m_3 > m_1 = m_2 = \frac{1}{3} - \epsilon$; Case 2: $m_3 < m_1 = m_2 = \frac{1}{3} + \epsilon$. For any non-uniform mass distribution, the orbits of the two equal masses m_1 (in blue) and m_2 (in red) are reflectional image of each other with respect to the symmetry axis of that of m_3 (in green) connecting its two lobes' mid-points. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

fact, in either limiting case, the associated shape curve is confined to the meridian corresponded to isosceles configurations of type 3 on the shape space M^*.

It is worth pointing out that the structure of our family of periodic orbits is in conformity with the *principle of natural termination*, obtained empirically by Strömgren [16,17] on the basis of numerical explorations and proved by Wintner [20] and Birkhoff [2]. The essence of the principle is as follows:

> Starting from any given orbit in a family, one can move along the family in two and only two directions in each of which the family has a *natural termination* (see Fig. 6).

One may continue in this fashion obtaining even more interesting families of periodic orbits, both in theory and in real solar systems (see Appendix A).

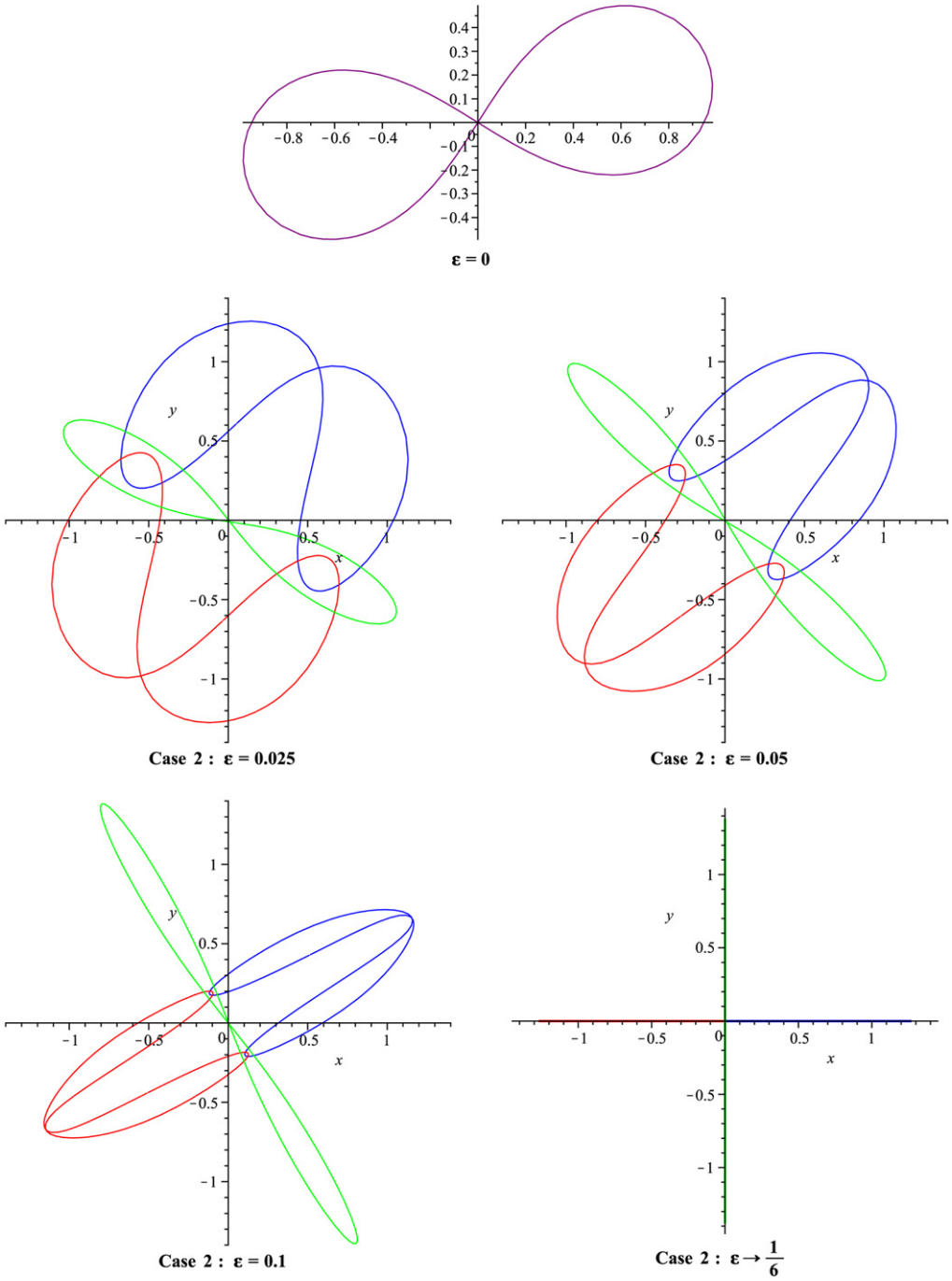

Fig. 1. (*continued*)

2. Foundations

The classical 3-body problem in celestial mechanics studies the local and global geometry of trajectories of a 3-body system as a conservative system with potential energy $-U$, where

$$U = \sum_{i<j} \frac{m_i m_j}{r_{ij}} \tag{1}$$

is the Newtonian potential function and $r_{ij} = |\mathbf{a}_i - \mathbf{a}_j|$ are the mutual distances. Here $m_i > 0$ are the masses, and $\mathbf{a}_i = (x_i, y_i, z_i)$ are the position vectors in Euclidean 3-space with respect to an inertial frame. Then the trajectories are locally characterized by Newton's equations

$$m_i \ddot{\mathbf{a}}_i = \frac{\partial U}{\partial \mathbf{a}_i} = \frac{m_i m_j}{r_{ij}^3}(\mathbf{a}_j - \mathbf{a}_i) + \frac{m_i m_k}{r_{ik}^3}(\mathbf{a}_k - \mathbf{a}_i), \quad \{i, j, k\} = \{1, 2, 3\}. \tag{2}$$

This equation is a 2-order ODE in Euclidean space $\mathbb{R}^9$, and hence a trajectory is completely determined by the initial positions and velocities of the particles – in agreement with the deterministic laws of classical mechanics.

The basic kinematic quantities are

$$I = \sum m_i |\mathbf{a}_i|^2, \qquad T = \frac{1}{2}\sum m_i |\dot{\mathbf{a}}_i|^2, \qquad \boldsymbol{\Omega} = \sum m_i (\mathbf{a}_i \times \dot{\mathbf{a}}_i) \tag{3}$$

which are, respectively, the total (polar) moment of inertia, kinetic energy, and the angular momentum. Their interactions with the potential function U play a major role in the dynamics of the 3-body problem. In fact, it is fairly easy to deduce the classical conservation laws of the system (2), namely the invariance of the linear momentum $\sum m_i \dot{\mathbf{a}}_i$, that of the angular momentum vector $\boldsymbol{\Omega}$, and of the total energy

$$h = T - U. \tag{4}$$

Moreover, by the t-derivative of I twice and using (2) we get the *Lagrange–Jacobi* equation:

$$\ddot{I} = 4T + 2\sum \mathbf{a}_i \cdot \frac{\partial U}{\partial \mathbf{a}_i} = 4T - 2U = 2(U + 2h) \tag{5}$$

which follows from Euler's theorem and the fact that U is homogeneous of degree -1 as a function of the position vectors $\mathbf{a}_i$.

On the other hand, the trajectories can be globally characterized using the basic action principles in mechanics due to Lagrange and Hamilton. These principles are quite different but somehow dual to each other. In either case, trajectories can be determined as solutions of a suitable boundary value problem – namely that for a given pair of points P, Q, what the trajectories

$$\gamma(t), \quad t_0 \leqslant t \leqslant t_1$$

with $\gamma(t_0) = P$ and $\gamma(t_1) = Q$ are. The solutions are extrema of an action integral $J(\gamma)$, on all virtual motions between the given pairs of points P, Q, of any of the following two types:

$$\textit{Lagrange}: \quad J_1(\gamma) = \int_\gamma T\,dt, \quad \textit{fixed energy } h; \tag{6}$$

$$\textit{Hamilton}: \quad J_2(\gamma) = \int_\gamma (T+U)\,dt, \quad \textit{fixed time interval } [t_0, t_1]. \tag{7}$$

Now, as usual, one can utilize the invariance of linear momentum by choosing the origin of the inertial frame at the center of mass. This reduces the study of 3-body trajectories to that of the associated time parametrized curves

$$\gamma(t) = \big(\mathbf{a}_1(t), \mathbf{a}_2(t), \mathbf{a}_3(t)\big)$$

in the 6-dimensional Euclidean (*general*) configuration space

$$M_{(g)} = \Big\{(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3);\ \sum m_i \mathbf{a}_i = 0\Big\} \cong \mathbb{R}^6. \tag{8}$$

The geometric reduction method that will be used in this project dates back to Jacobi (1840), who geometrized classical mechanics by reformulating Lagrange's least action principle. It is in fact worth noting that in the action integral $J_1(\gamma)$, as in (6), time is allowed to vary; i.e., the limit of integration is not fixed. This awkwardness led Jacobi to suggest that the time differential be eliminated from $J_1(\gamma)$. He introduced the *kinematics metric* on $M_{(g)}$

$$ds^2 = 2T\,dt^2 = \sum m_i\big(dx_i^2 + dy_i^2 + dz_i^2\big) \tag{9}$$

which represents the kinetic energy. Then, for a fixed energy level h, he considered the modified *dynamical* metric

$$ds_h^2 = (U+h)\,ds^2 \tag{10}$$

and observed that

$$\sqrt{2}J_1(\gamma) = \sqrt{2}\int_\gamma T\,dt = \int_\gamma \sqrt{U+h}\,ds = \int_\gamma ds_h$$

is the arc-length of the virtual motion γ in $M_{(g)}$ with the Riemannian metric (10). Consequently, trajectories of Newton's equations at a fixed energy level h are precisely the geodesics in $M_{(g)}$ with respect to the metric ds_h^2 (for further information on this geometric approach, see [1,8,9]).

2.1. $SO(2)$-symmetry and reduction to the congruence moduli and shape space level

In order to contribute to Hsiang–Straume's setting, we shall assume the following setup. A three-body motion with vanishing angular momentum is always confined to a fixed plane (for purely kinematic reasons), so the motions we shall study are always planar. Therefore, we choose a plane $\mathbb{R}^2 \subset \mathbb{R}^3$ with unit normal vectors $\pm\mathbf{n}$ and define an *oriented m-triangle* to be a pair $(\mathbf{X}, \mathbf{n})$, where $\mathbf{X} = (\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ represents the position of the 3-body system in the plane $\mathbb{R}^2$ (i.e., $\mathbf{a}_i \cdot \mathbf{n} = 0$ for all i) constrained by the center of mass condition in (8). An m-triangle is called *degenerate* if the three masses are aligned. A non-degenerate m-triangle is said to be *positively* (resp. *negatively*) *oriented* if $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{n})$ is a right-handed (resp. left-handed) frame at every instant of time. The squared norm of

X with respect to the kinematic metric is the moment of inertia $|\mathbf{X}|^2 = I = \rho^2$, so ρ is the natural size function for m-triangles. For our purpose, we shall modify the definition (8) of the (general) configuration space by taking the subspace

$$M \cong \mathbb{R}^4 \subset \mathbb{R}^6 : \sum_{i=1}^{3} m_i \mathbf{a}_i = 0 \tag{11}$$

as our *configuration space* which consists of the above m-triangles in the fixed plane $\mathbb{R}^2$. The rotation group $SO(2)$ acts naturally on m-triangles, and the $SO(2)$-orbit of an m-triangle is its congruence class in the usual geometric sense.

It is convenient to replace the Euclidean space M in (11) by the space of all oriented m-triangles having the obvious induced action of $SO(2)$ and an invariant kinematic Riemannian structure. Then, there is the map projection

$$\pi : M \to \bar{M} = M/SO(2) \cong \mathbb{R}^3 \tag{12}$$

which identifies the orbit space $\bar{M}$, called the *congruence moduli space*, with the usual 3-space. This fits in such a fashion that the equatorial plane $z = 0$ represents congruence classes of degenerate triangles, and the semi-space $z > 0$ (resp. $z < 0$) represents positively (resp. negatively) oriented m-triangles. It turns out that $\bar{M}$ is homeomorphic to $\mathbb{R}^3$, as indicated in (12), and they are also diffeomorphic away from the origin ($\rho = 0$). Naturally, the subset $M^* = (\rho = 1)$ or *unit sphere* of $\bar{M}$ represents similarity classes of m-triangles and is, therefore, called the *shape space*. Namely, a point in M^* represents a homothety class of a triangle (of size $\rho > 0$), and it is an important fact that the shape space is actually the 2-sphere

$$M^* \cong S^2 : x^2 + y^2 + z^2 = \rho^4 = 1. \tag{13}$$

However, with the induced metric $d\sigma^2 = d\bar{s}^2|_{M^*}$, the shape space is actually a round sphere of radius 1/2

$$\big(M^*, d\sigma^2\big) = S^2(1/2).$$

Therefore, as a Riemannian cone over M^*, the kinematic metric on the moduli space $\bar{M}$ can be expressed as

$$d\bar{s}^2 = d\rho^2 + \rho^2\, d\sigma^2 = d\rho^2 + \rho^2 \frac{(d\varphi^2 + \sin^2\varphi\, d\theta^2)}{4}, \tag{14}$$

where (ρ, φ, θ) are the spherical coordinates on $\bar{M} \cong \mathbb{R}^3$ and, (φ, θ) is any choice of the spherical polar coordinates on $M^* \cong S^2$.

A motion of m-triangles is a parametrized curve $t \to \gamma(t)$ in the configuration space M. Such a curve may, for example, be a solution of Newton's equations and hence represent a solution (or trajectory) of the 3-body problem. The above reduction technique can replace $\gamma(t)$ by either its moduli curve or shape curve

$$\bar{\gamma}(t) = \big(\rho(t), \varphi(t), \theta(t)\big), \qquad \gamma^*(t) = \big(\varphi(t), \theta(t)\big) \tag{15}$$

respectively in the 3-dimensional space $\bar{M} = \mathbb{R}^3$ and its 2-sphere $M^* = S^2$. The kinetic energy $\bar{T}$ of $\bar{\gamma}(t)$ is encoded by the above metric (14), namely

$$d\bar{s}^2 = 2\bar{T}\,dt^2, \qquad \bar{T} = T - T^{\omega}, \tag{16}$$

where T^{ω} is the purely rotational kinetic energy of the motion $\gamma(t)$, which can be determined explicitly from $\gamma(t)$ and the angular momentum vector $\boldsymbol{\Omega}$. The reconstruction of the motion $\gamma(t)$ from the knowledge of the curve $\bar{\gamma}(t)$, with a given constant angular momentum vector $\boldsymbol{\Omega}$, is a purely mechanical lifting procedure in (12), which yields a unique curve $\gamma(t)$ up to congruence. By (14) and (16), the total kinetic energy can be written (cf. [9, Ch. 2]) as

$$T = \bar{T} + T^{\omega} = \frac{1}{2}\dot{\rho}^2 + \frac{\rho^2}{8}\big(\dot{\varphi}^2 + \big(\sin^2\varphi\big)\dot{\theta}^2\big) + T^{\omega} \tag{17}$$

where the rotational term T^{ω} vanishes precisely when $\boldsymbol{\Omega} = 0$ which is our standing assumption in the study of three-body motions.

2.2. Dynamics in the moduli space

We have seen that the study of 3-body motions can be reduced to that of their corresponding moduli curves in $\bar{M}$ which are the solutions of appropriate reduced differential equations (of type Newton, Lagrange, Jacobi, or Hamilton), or those of some least action principle at the level of $\bar{M}$. The homogeneity of the Newtonian potential function (1) allows us to write

$$U(\rho, \varphi, \theta) = \frac{1}{\rho}U^*(\varphi, \theta)$$

where U^* is the *shape potential* function on the shape space $M^* = S^2$.

To the special case of planar 3-body motions, there corresponds the following system of ODEs in $\bar{M}$ (cf. [15, (5.11)]):

$$\begin{aligned}
&\text{(i)} \quad 0 = \ddot{\rho} + \frac{\dot{\rho}^2}{\rho} - \frac{1}{\rho}\left(\frac{1}{\rho}U^* + 2h\right),\\
&\text{(ii)} \quad 0 = \ddot{\varphi} + 2\frac{\dot{\rho}}{\rho}\dot{\varphi} - \frac{1}{2}\sin(2\varphi)\dot{\theta}^2 - \frac{4}{\rho^3}U^*_{\varphi},\\
&\text{(iii)} \quad 0 = \ddot{\theta} + 2\frac{\dot{\rho}}{\rho}\dot{\theta} + 2\cot(\varphi)\dot{\varphi}\dot{\theta} - \frac{4}{\rho^3}\frac{1}{\sin^2(\varphi)}U^*_{\theta},
\end{aligned} \tag{18}$$

where Eq. (i) is just the Lagrange–Jacobi equation (5).

3. The remarkable family of periodic three-body orbits

Let (m_1, m_2, m_3) be any given mass distribution with $m_1 = m_2$ in the *mass space*

$$\big\{(m_1, m_2, m_3) \in \mathbb{R}^3 \;\big|\; m_i > 0,\ m_1 + m_2 + m_3 = 1\big\}.$$

We call a mass distribution *uniform* if $m_1 = m_2 = m_3$, and hence it is called non-uniform if at least one of the three masses is different from the others.

Let $T > T^*$ be any positive real numbers. Assuming σ to be the generator of the two-element dihedral group $D_1 = \mathbb{Z}_2$, we can define the actions of $\mathbb{Z}_2$ on $\mathbb{R}/T\mathbb{Z}$ and $\mathbb{R}^2$ as follows

$$\sigma \cdot t = -t + T^*, \qquad \sigma \cdot \boldsymbol{v} = R_{\ell}(\boldsymbol{v}),$$

where R_{ℓ} stands for the reflection in the line ℓ in which $\boldsymbol{v}$ being reflected; the group generator may be written as $\sigma = \sigma_{(T^*,\ell)}$.

Theorem. *Let (m_1, m_2, m_3) be any given mass distribution with $m_1 = m_2$.*

(A) There exist three "deformed" 8-shaped planar loops $\boldsymbol{b}_i : \mathbb{R}/T\mathbb{Z} \to \mathbb{R}^2$, $i = 1, 2, 3$, for some $T > 0$, with the following properties:

(i) *for each t, $\sum_1^3 m_i \boldsymbol{b}_i(t) = 0$;*
(ii) *$\boldsymbol{b}_1$, $\boldsymbol{b}_2$ are conjugate and $\boldsymbol{b}_3$ is equivariant, both under the actions of the dihedral group $D_1 = \mathbb{Z}_2$ with generator $\sigma = \sigma_{(T^*,\ell)}$, for some $0 < T^* < T$, as being defined above*:

$$\sigma \cdot \big(\boldsymbol{b}_3(t)\big) = \boldsymbol{b}_3(\sigma \cdot t), \qquad \sigma \cdot \big(\boldsymbol{b}_i(t)\big) = \boldsymbol{b}_j(\sigma \cdot t), \quad \{i, j\} = \{1, 2\}$$

where ℓ is the symmetry axis of $\boldsymbol{b}_3$ connecting its two lobes' mid-points, and the lobes of $\boldsymbol{b}_3$ can be given by $\boldsymbol{b}_3([0, T^*])$ and $\boldsymbol{b}_3([T^*, T])$;
(iii) *the orbit $\gamma : \mathbb{R}/T\mathbb{Z} \to M$ defined by*

$$\gamma(t) = \big(\boldsymbol{b}_1(t), \boldsymbol{b}_2(t), \boldsymbol{b}_3(t)\big)$$

is a vanishing angular momentum T-periodic solution of the three-body problem, where $\boldsymbol{b}_i = \mathbf{a}_i$, $i = 1, 2, 3$, are the position vectors of the bodies.

(B) Moving along the line segment $m_1 = m_2$ on the mass space in the two directions with respect to (the heavier or lighter) m_3, yields two continuous families of periodic three-body motions with non-uniform mass distributions of the form γ as in $(A)_{\text{(iii)}}$ which connect at the celebrated figure-8 orbit E in the equal-mass case, and yield a continuous family of vanishing angular momentum periodic solutions of the three-body problem in the plane (see Fig. 1).

Remark. At one end of the family as introduced in (B) above, when the two equal masses are infinitesimal and the third one reaches the value of $+1$, the equal masses describe an elliptic Keplerian motion on a circle centered at the origin, but in opposite directions around m_3 which rests at the origin (i.e., we have a double Kepler problem). In fact, the infinitesimal equal masses move on the two different semicircles symmetric with respect to the y-axis. At the other end of the family, when the third mass is infinitesimal, orbits of the bodies are flattened out so that the equal masses move along the x-axis and m_3 moves along the y-axis (i.e., we have a special case of periodic solutions of a restricted 3-body problem).

The remainder of the paper is devoted to a proof of this theorem.

4. Computer experiments

Every period in the history has its own version of what the *three-body problem* is. For Weierstrass, Poincaré and Sundman (cf. [18,19]), about a century ago, the basic question was how to find a convergent power series expansion along the orbit. However, nowadays, we can in fact *solve* any given three-body problem, starting at given initial positions and velocities of the three masses, by calculating the resulting orbit using a computer.

Although this is certainly of crucial importance in modern space science and technology, here our main aim of computing via a Computer Algebra System (CAS) is to get insight into a deeper understanding of the qualitative behavior of the periodic three body problems having two equal masses slightly different from the third one, and not just the numbers and limited amount of information which one can derive from orbit calculations. Our computer algebra work-sheets in this project are based on the `Maple 12` software system, one of the most powerful CASs currently available.

As stated in Section 2, for planar three-body motions of a given constant total energy which has been fixed to $h = -0.5$ in our computations, the corresponding congruence moduli curves $\bar{\gamma}(t)$ are the solutions of the system (18) at the level of $\bar{M}$.

In this section, we shall solve the system (18) numerically with various systematically chosen initial values in the spherical coordinates (ρ, φ, θ), where (φ, θ) denotes a choice of spherical polar coordinates on the shape sphere $M^* = S^2(1)$, in order to find an approximate solution curve of the shape curve in the Theorem. Here and subsequently, we focus our attention to the geometric study of the shape curve

$$\gamma^*(t) = (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)(t)$$

and to that of the Newtonian shape potential function $U^* = U^*(\varphi, \theta)$.

For convenience we set the binary collision point b_{23} into the point corresponding to the spherical polar coordinates $(\varphi, \theta) = (\frac{\pi}{2}, 0)$ on M^*, hence

$$b_{12} = (\cos\beta_2, -\sin\beta_2, 0), \qquad b_{23} = (1, 0, 0), \qquad b_{31} = (\cos\beta_3, \sin\beta_3, 0)$$

where $\beta_i = \arccos(\frac{m_1 m_j - m_i}{m_1 m_j + m_i}) \in (0, \pi)$ is the longitude angle of the binary collision point b_{i1}, $\{i, j\} = \{2, 3\}$. Without loss of generality we can tacitly assume that a fixed mass distribution (m_1, m_2, m_3) with $m_1 = m_2$ is given, and it is normalized so that $m_1 + m_2 + m_3 = 1$. Accordingly, see [9],

$$U^* = \sum_{i=1}^{3} \frac{\hat{m}_i^{3/2} (m_i^*)^{-1/2}}{\sqrt{1 - \sin\varphi\cos(\theta - \theta_i)}}, \tag{19}$$

where $(\theta_1, \theta_2, \theta_3) = (0, \beta_3, -\beta_2)$, $\hat{m}_i = m_j m_k$, and $m_i^* = \frac{1}{2}(1 - m_i)$, $\{i, j, k\} = \{1, 2, 3\}$, the latter gives the dual mass distribution with $\sum m_i^* = 1$.

By the existence and uniqueness theorem of a system of ODEs, we only need suitable initial values very close to the actual ones to run our program; but, who knows where in the continuous forest of real numbers the actual ones live.

At the beginning of the second millennium, a remarkable periodic orbit for the planar three-body problem with the uniform mass distribution, i.e. with $m_i = 1/3$, was exhibited by A. Chenciner and R. Montgomery using a variational method in [3]. The orbit known as the "figure-eight" has zero angular momentum and a very rich symmetry pattern. In addition, it is worth noting that the numerical computations by Carles Simó illuminated the way of exploration in [3]; e.g., the computations indicated that the figure-eight orbit is a completely elliptic one with torsion (cf. [3,13]).

4.1. Initial values estimations

Here, we shall make use of the *Hopf fibration map* as a practical tool to transfer the required data from the level of M to that of $\bar{M}$ where our system of ODEs (18) lives. We begin by introducing the explicit formulas for the standard Jacobi vectors $(\mathbf{x}_1, \mathbf{x}_2)$ which, following [14], can be expressed as

$$\mathbf{x}_1 = \sqrt{\frac{m_1}{m_2 + m_3}}\mathbf{a}_1, \qquad \mathbf{x}_2 = \sqrt{\frac{m_2(m_2 + m_3)}{m_3}}\left(\mathbf{a}_2 + \frac{m_1}{m_2 + m_3}\mathbf{a}_1\right).$$

It is easily seen that

$$|\mathbf{x}_1|^2 + |\mathbf{x}_2|^2 = \sum_i m_i |\mathbf{a}_i|^2 = I\left(= \rho^2\right). \tag{20}$$

Recall that in the case study of planar three-body motions, due to the conservation of center of mass with respect to time, the configuration space $M = \mathbb{R}^4$ and for the moduli space we have the identification $\bar{M} = \mathbb{R}^4/SO(2) \cong \mathbb{R}^3$, cf. (12). Letting $S^3(\rho)$ and $S^2(\rho^2)$ be the round spheres respectively of radius ρ and ρ^2 and using our metric setting (cf. Section 2), the Hopf fibration (orbit) map reads

$$H: S^3(\rho) \subset M \to S^2\big(\rho^2\big) \subset \mathbb{R}^3 \cong \bar{M},$$
$$(\mathbf{x}_1, \mathbf{x}_2) \to (x, y, z)$$

where in terms of the standard inner- and cross-product of Euclidean 3-space with $\mathbf{k}$ as the third standard basis vector

$$x = |\mathbf{x}_1|^2 - |\mathbf{x}_2|^2, \qquad y = 2\langle \mathbf{x}_1, \mathbf{x}_2 \rangle, \qquad z = 2\langle \mathbf{x}_1 \times \mathbf{x}_2, \mathbf{k} \rangle. \tag{21}$$

Note that since the motion is planar, the vectors $\mathbf{x}_1 \times \mathbf{x}_2$ and $\mathbf{k}$ are collinear. Retrieving the Cartesian coordinates $(x, y, z) \in S^2(\rho^2)$ from the spherical ones, we can see

$$x^2 + y^2 + z^2 = \big(|\mathbf{x}_1|^2 + |\mathbf{x}_2|^2\big)^2 = \rho^4.$$

By (20) and (21), $x + \rho^2 = 2|\mathbf{x}_1|^2$ or equivalently

$$\rho^2(1 + \sin\varphi\cos\theta) = 2\left(\frac{m_1}{1 - m_1}\right)|\mathbf{a}_1|^2.$$

Differentiating both sides with respect to time, gives

$$2\rho\dot{\rho}(1 + \sin\varphi\cos\theta) + \rho^2(\dot{\varphi}\cos\varphi\cos\theta - \dot{\theta}\sin\varphi\sin\theta) = 4\left(\frac{m_1}{1 - m_1}\right)\langle \mathbf{a}_1, \dot{\mathbf{a}}_1 \rangle. \tag{22}$$

In addition, by the conformal modification of the standard Euclidean metric, the kinetic energy T can be written as

$$T = \frac{1}{2}\dot{\rho}^2 + \frac{\rho^2}{8}\big(\dot{\varphi}^2 + \big(\sin^2\varphi\big)\dot{\theta}^2\big). \tag{23}$$

In what follows, the *canonical symmetry plane* plays a key role.

Definition 4.1. By the *canonical symmetry plane* we mean the plane sitting in the moduli space $\bar{M} = \mathbb{R}^3$ in which both the z-axis and the line passing through the Euler point $e_3 = (\cos(\beta_3/2), \sin(\beta_3/2), 0)$ and its antipodal point (i.e., the binary collision point b_{12}) lie, see Fig. 2.

Actually, here, we are interested in finding a solution shape curve on the shape space $M^* = S^2(1)$ which meets the symmetry plane orthogonally.

We are now in a position to estimate the initial values. For any suitably given $\epsilon > 0$, let

$$m_1 = m_2 = \frac{1}{3} \mp \epsilon, \qquad m_3 = \frac{1}{3} \pm 2\epsilon$$

be the normalized mass distribution for our m-triangles. We have divided the numerical computations into two cases whose procedures are almost the same.

`Case 1`: $m_3 > m_1 = m_2$.

To make beneficial use of Carles Simó's numerical computations on his E orbit (see for instance [13]), in the case of uniform mass distribution, we should begin our exploration by choosing a sufficiently small $\epsilon > 0$, say $\epsilon = 10^{-6}$, in order to find the suitable initial values. As, for such small values of ϵ, we can estimate the suitable initial values just by perturbing those of Simó's E orbit little by little. If a three-body motion with uniform mass distribution satisfies the following conditions

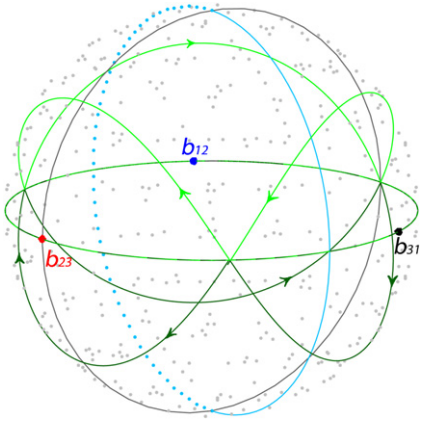

Fig. 2. Shape curve on the shape space. The shape point b_{ij} represents the binary collision of m_i and m_j.

$$\Omega = 0, \qquad \frac{dI}{dt}(0) = 0, \qquad \mathbf{a}_3(0) = 0$$

then straightforward calculations give

$$\begin{aligned} &\mathbf{a}_1 = -\mathbf{a}_2, \qquad \mathbf{a}_3 = 0, \\ &\dot{\mathbf{a}}_1 = \dot{\mathbf{a}}_2 = \frac{-1}{2}\left(\frac{m_3}{m_1}\right)\dot{\mathbf{a}}_3 \end{aligned} \tag{24}$$

which lead to the initial values of Simó's E orbit. Note that, as soon as $\dot{\mathbf{a}}_3$ is known, the initial positions of the bodies in an Euler configuration for the E orbit can be recovered by applying the energy, angular momentum and center of mass integrals.

Now, let us turn to our estimation method. Let $\xi \in \mathbb{R}$ be the perturbation correction in the third body's initial velocity

$$\dot{\mathbf{a}}_3^{\epsilon} := \dot{\mathbf{a}}_3^0 - \xi$$

where the superscript $\epsilon \geqslant 0$ refers to our ϵ-dependent mass distribution. For convenience, from now on, $\dot{\mathbf{a}}_3$ (ignoring the superscript) stands for the initial velocity of the third body after applying the perturbation correction for any given $\epsilon > 0$.

By (24) and using the afore-mentioned perturbation corrections, the kinetic energy at $t = 0$ reads

$$T(0) = \frac{1}{4}\left(\frac{m_3}{m_1}\right)|\dot{\mathbf{a}}_3|^2. \tag{25}$$

Substituting (19) and (25) into the energy integral $h = -0.5 = T - \frac{U^*}{\rho}$, we can see that at $t = 0$

$$\rho(0) = \left(\frac{2(1-m_3)}{m_3(|\dot{\mathbf{a}}_3|^2 - 1) + 1}\right)U^*(\varphi, \theta). \tag{26}$$

In addition, by (23) and (25)

$$\dot{\varphi}^2 + (\sin^2\varphi)\dot{\theta}^2 = \frac{2}{\rho^2}\left[\left(\frac{m_3}{m_1}\right)|\dot{\mathbf{a}}_3|^2 - 2\dot{\rho}^2\right], \quad \text{at } t = 0. \tag{27}$$

Let us temporarily set $\epsilon = 10^{-6}$, and

$$\varphi(0) = \frac{\pi}{2}, \qquad \theta(0) = \frac{\beta_3}{2}, \qquad \dot{\rho}(0) = 0 \tag{28}$$

which are in conformity with the initial conditions of Simó's E orbit (cf. [13]). Therefore, we are left with the task of estimating the rest of initial values, namely $\rho(0)$, $\dot{\varphi}(0)$ and $\dot{\theta}(0)$.

By (19), (26) and (28), the initial size $\rho(0)$ of our m-triangle can be estimated immediately.

Using the fact that the initial positions (namely $\mathbf{a}_i$'s) can be recovered as soon as $\dot{\mathbf{a}}_3$ is known and by (28), the system of equations formed by (22) and (27) is a system of two equations in two variables $\dot{\varphi}(0)$, $\dot{\theta}(0)$ which gives two pairs of solutions. Experimentally speaking, we should choose the pair in which $\dot{\varphi}(0) > 0$.

It remains to check how suitable the estimated initial values are, and with which accuracy the solution shape curve meets the symmetry plane orthogonally. For this purpose, we define t_m to be the time at which the solution shape curve meets the symmetry plane for the first time, and η to be the acute angle between the velocity vector $\dot{\gamma}^*(t_m)$ and a unit normal vector of the symmetry plane. Needless to say that the smaller value the η takes, the more desirable solution the solution shape curve is.

For the sake of convenience, an explicit formula for η is

$$\eta = \arccos\left(\frac{\langle \dot{\gamma}^*(t_m), \vec{n}\rangle}{\nu(t_m)}\right), \tag{29}$$

where $\nu(t) = |\dot{\gamma}^*(t)| = \sqrt{\dot{\varphi}^2(t) + (\sin^2\varphi(t))\dot{\theta}^2(t)}$, and $\vec{n} = (\sin\beta_2, \cos\beta_2, 0)$ (or else its supplementary angle in order to have the acute angle η being uniquely defined). By Definition 4.1, the first meeting time t_m must satisfy

$$\theta(t_m) + \beta_2 = 0 = \theta(t_m) + \pi - \theta(0). \tag{30}$$

Applying our method of perturbation corrections repeatedly for the temporarily chosen $\epsilon = 10^{-6}$ shows that the angle η can tend to zero for suitable sets of the initial values. Geometrically speaking, our method gives more, namely smoothness of the solution shape curve and that the shape curve returns to its initial position two times with a high accuracy. However, the plots can still be distinguished from the desired ones.

Case 2: $m_3 < m_1 = m_2$.

In this case, our m-triangles have the following ϵ-dependent mass distribution

$$m_1 = m_2 = \frac{1}{3} + \frac{\epsilon}{2}, \qquad m_3 = \frac{1}{3} - \epsilon$$

which is in fact the dual mass distribution of the one in Case 1 for $0 < \epsilon \leqslant \frac{1}{3}$.

In the same manner as in Case 1 we can see that almost the same conclusions can be drawn for the solution shape curve.

Let us temporarily set $\epsilon = 10^{-6}$ and $\Delta\rho = \rho^{\epsilon}(0) - \rho^{0}(0)$, where

$$\rho^{\epsilon}(0) \approx 0.233543784396839624536326744,$$

$$\rho^{0}(0) \approx 0.233543786355978198307429133$$

are respectively the initial size of our ϵ-dependent m-triangle and that of Simó's E orbit [13] estimated by the same method as in Case 1.

We were surprised at finding out that the initial size of our ϵ-dependent m-triangle in Case 2 (i.e., when m_3 is lighter than the other two equal masses) is smaller than that of our m-triangle in Case 1, but their corresponding $\Delta\rho$'s have approximately the same absolute values.

Note that, in order to get a desired solution shape curve γ^* (namely the one which meets the symmetry plane orthogonally at the local extrema of its *m-latitude function* $\lambda_{\gamma^*}(t)$ $(= \frac{\pi}{2} - \varphi(t))$, is smooth and closes up with a reasonably high accuracy) for slightly bigger $\epsilon > 0$ than our temporarily chosen $\epsilon = 10^{-6}$ in both Cases 1 and 2, in addition to applying our method of perturbation corrections, we need to deform the initial configuration of our m-triangle as well; e.g., by some modifications in $\varphi(0)$, and fixing $\theta(0) = \frac{\beta_3}{2}$, $\dot\rho(0) = 0$ (which will be described in detail in Section 5).

Remark 4.2. Repeated application of the afore-mentioned method of perturbation corrections together with some modifications in the initial values of the system (18) enables us to conjecture that

- the solution shape curve is periodic,
- meets the symmetry plane orthogonally at $t \in \{nt_m + (n-1)t'_m \mid n \in \mathbb{N}\}$ at which the m-latitude function takes its local extrema, and
- returns to its initial position at $t = 2t_m,\ 2t_m + 2t'_m, \ldots$

for any ϵ-dependent mass distribution, where $0 < \epsilon \leqslant \frac{1}{3i}$ in case $i = 1, 2$.

5. Analytical proof

Recall that for any suitably given $\epsilon > 0$ as in Remark 4.2, our normalized ϵ-dependent mass distribution is given by

$$m_1 = m_2 = \frac{1}{3} \mp \epsilon, \qquad m_3 = \frac{1}{3} \pm 2\epsilon.$$

As stated in Section 4, in order to estimate the actual initial values for bigger ϵ's (other than applying our method of perturbation corrections) we have needed to make some modifications in $\varphi(0)$. Hence, in contrast to Simó's E orbit, for our three-body motions the associated solution shape curves on the shape space $M^* = S^2(1)$ cannot start at the Euler point e_3 anymore. Accordingly, the best we can hope for is that our shape curves keep the symmetry with respect to the canonical symmetry plane.

5.1. The cornerstones of the proof

The proof is based on Hsiang–Straume's *unique parametrization* and *monotonicity theorems* (cf. respectively Theorems 4.6 and 5.8 in [8]) which can be stated as follows.

Lemma 5.1 *(unique parametrization). A three-body motion with vanishing angular momentum is uniquely determined by its oriented geometric (i.e., arc-length parameterized) shape curve γ^* on the 2-sphere M^*.*

Formulating the monotonicity theorem, which is also referred to as the *monotone m-latitude theorem*, needs some relevant technical material from [8]; to make our exposition self-contained, we repeat these materials here.

For a given ϵ-dependent mass distribution (m_1, m_2, m_3), let the points $\mathbf{p}_0, \hat{\mathbf{p}}_0 \in M^*$ represent the pair of equilateral m-triangles with opposite orientations which are actually the minima of U^*. For a uniform mass distribution the pair of minima of U^* happens to coincide with the North and South poles of M^*, namely $\{N, S\}$, but this does not hold for the case of non-equal masses. However, there exists a unique Möbius transformation which maps $\mathbf{p}_0$ to N, $\hat{\mathbf{p}}_0$ to S and does the equator circle to itself.

For a given (smooth) shape curve $\gamma^*(t)$, the m-latitude function is defined by

$$\lambda_{\gamma^*}(t) = \lambda\big(\gamma^*(t)\big) = \frac{\pi}{2} - \varphi(t) \tag{31}$$

which records the m-modified latitude, under the afore-mentioned Möbius transformation, along the curve.

Lemma 5.2 *(monotonicity). Let $\gamma^*(t)$ be the associated shape curve of a three-body motion with vanishing angular momentum, and λ_{γ^*} be the m-latitude function as in* (31)*. Then λ_{γ^*} is a strictly monotonic function along γ^* between every two succeeding local maxima and minima which lie on opposite hemispheres of the* 2*-sphere M^*. In particular, the shape curve intersects the eclipse circle* (*equator*) *at a unique point between every two such local extrema* (*see* [6] *for a similar type of monotonicity theorem*)*.*

The main idea of the proof is to make a practical use of the *method of analytic continuation* which guarantees the existence and uniqueness of our desired 4π-periodic shape curves associated to the periodic solutions as in the Theorem.

The proof will be divided into three steps.

Step 1. Applying the method of analytic continuation guarantees that for any given ϵ-dependent mass distribution there exists a unique set of initial values leading to a solution shape curve which meets the symmetry plane orthogonally at the local extrema of its m-latitude function.

Step 2. Lemma 5.2 (which describes a type of piecewise monotonic behavior of the shape curve γ^*) shows that the shape curve is smooth and 4π-periodic in time $T = 2(t_m + t'_m)$; it returns to its initial position two times at $t = 2t_m$, T, where t_m is the first orthogonally meeting time as in (30) and $2t_m + t'_m$ is the second one (cf. Remark 4.2). This follows by reflecting the arcs of our shape curve obtained up to each meeting time about the symmetry plane, and from the fact that the oriented shape curve is geometrically invariant under time translations, $t \to t + t_0$, as well as time reversal, $t \to -t$.

Step 3. The only point remaining concerns the behavior of the three-body motion $\gamma(t)$ in the configuration space M. Since $\gamma(t)$ can be determined up to congruence by its moduli curve $\bar{\gamma}(t)$ in $\bar{M}$, Lemma 5.1 shows that $\gamma(t)$ can be uniquely determined, up to congruence, by its oriented geometric shape curve γ^* on the 2-sphere M^*. In particular, by the kinematic Gauss–Bonnet theorem [9, Th. C2] and the periodicity of the shape curve in time $T = 2(t_m + t'_m)$ (cf. (30)), the motion $\gamma(t)$ is periodic. Using a simple formula for the *individual torques* of the three bodies, we prove that the third body m_3 moves on an 8-shaped orbit (called $\boldsymbol{b}_3$), and the two equal masses m_1, m_2 move on different deformed ones (resp. called $\boldsymbol{b}_1$, $\boldsymbol{b}_2$) which are reflectional image of each other with respect to the symmetry axis of the orbit b_3 connecting its two lobes' mid-points.

The remainder of this section will be devoted to the analytical proof of the Theorem whose quick sketch was just given.

For any given $t_0 \geqslant 0$, it is evident that a Θ-parametrization of the latitude circle of the 2-sphere (hit by the shape curve at $t = t_0$) projected on the equatorial plane can be represented by

$$E_{t_0} := \Theta \mapsto \big(\sin\varphi(t_0)\cos\Theta, \sin\varphi(t_0)\sin\Theta, 0\big),$$

and that $\frac{dE_{t_0}(\theta(t))}{dt}(t_0)$ is in fact a normal vector of the hyperplane $\theta = \theta(t_0)$ in $\bar{M}$.

Definition 5.3. The *inclination* of the shape curve at $t = t_0$, denoted by $\mu(t_0)$, is defined to be the acute angle between the velocity vector $\dot{\gamma}^*(t_0)$ and $\frac{dE_{t_0}(\theta(t))}{dt}(t_0)$ which reads

$$\mu(t_0) = \arccos\left(\frac{|\dot{\theta}|\sin\varphi}{\sqrt{\dot{\varphi}^2 + \dot{\theta}^2\sin^2\varphi}}\right)(t_0).$$

Note that, since latitude circles on the 2-sphere M^* meet the symmetry plane orthogonally, the following conditions are equivalent (cf. (29), (30)):

$$\eta = 0 \quad \Longleftrightarrow \quad \mu(t_m) = 0. \tag{32}$$

Due to the symmetry with respect to the canonical symmetry plane, the shape curve meets the symmetry plane orthogonally at $t = t^*$ if and only if the m-latitude function takes a local extremum at the same meeting time $t^* \in \{t_m, 2t_m + t'_m, \ldots\}$. In order to check whether the shape curve meets the symmetry plane orthogonally, we have to check whether the moduli curve $\bar{\gamma}$ does so since both our system of ODEs (18), whose solution is $\bar{\gamma}$, and the canonical symmetry plane live in $\bar{M}$; moreover, we require the *structural arcs* of γ^* between each two consecutive orthogonally meeting times of the shape curve with the canonical symmetry plane to be congruent.

By Lemma 5.2, the shape curve γ^* can be viewed as a union of its segments between each two consecutive extrema of the m-latitude function which will be referred to as the *fundamental segments*. More precisely, a fundamental segment is of the form $\gamma^*|_{[t_0,t_1]}$, where $\dot{\varphi}(t_0) = 0 = \dot{\varphi}(t_1)$; while a structural arc is given by $\gamma^*|_{[t_{m_0},t_{m_1}]}$, where both $\dot{\rho}$, $\dot{\varphi}$ vanish at $t = t_{m_0}, t_{m_1}$. It is worth noting that each structural arc of a periodic shape curve can be constructed by connecting a finite number of such fundamental segments.

The following lemma is the key to constructing the fundamental segments and structural arcs of the solution shape curves.

Lemma 5.4. *The moduli curve $\bar{\gamma}(t)$ associated to a three-body trajectory $\gamma(t)$ meets the canonical symmetry plane orthogonally at $t = t^* \in \{t_m, 2t_m + t'_m, \ldots\}$ if and only if $\dot{\varphi}(t^*) = 0 = \dot{\rho}(t^*)$.*

Remark 5.5. The above lemma in fact holds if and only if

- $\mu(t^*) = 0$ (or equivalently $\dot{\varphi}(t^*) = 0$, cf. Definition 5.3), and
- $\langle \frac{d\bar{\gamma}}{dt}(t^*), \mathbf{b}_{12}\rangle$ vanishes,

where $\mathbf{b}_{12}$ is the binary collision point (cf. Fig. 2).

5.2. The analytic continuation method

For a given ϵ-dependent mass distribution, temporarily set

$$\delta := d\big[\big(\rho^0, \mu^0\big)(0), \big(\rho^\epsilon, \mu^\epsilon\big)(0)\big] + \epsilon, \tag{33}$$

where d stands for the Euclidean distance in the (ρ, μ)-plane $= \mathbb{R}^2 \subset T_{\bar{\gamma}(0)}\bar{M}$; $\rho^0(0)$, $\mu^0(0)$ are respectively the initial size of the m-triangle and the initial inclination of the associated shape curve of Simó's E orbit, and $\rho^\epsilon(0)$, $\mu^\epsilon(0)$ are those (estimated ones) of our three-body motion for the given $\epsilon > 0$ (cf. Section 4.1).

As stated at the beginning of Section 5, we need some modifications in $\varphi(0)$. Let $\Phi \in (0, \frac{\pi}{2})$ be the *φ-modification parameter* so that

$$\varphi(0) = \frac{\pi}{2} \mp \Phi \tag{34}$$

with "−" and "+" for the mass distributions in Case 1 and 2, respectively.

We are now in a position to take our first step in proving the Theorem (cf. Section 5.1). Let us set

$$\begin{aligned} \rho^{\epsilon,\delta}(0) &= \delta\cos\omega + \rho^0(0), \\ \tilde{\mu}^{\epsilon,\delta}(0) &= \delta\sin\omega + \mu^0(0), \end{aligned} \tag{35}$$

where ω varies on $[0, 2\pi)$, and $\tilde{\mu}^{\epsilon,\delta}(0)$ will be referred to as the *virtual initial inclination*. Clearly, the locus of the points $\{(\rho^{\epsilon,\delta}, \tilde{\mu}^{\epsilon,\delta})(0);\ \omega \in [0, 2\pi)\}$ is a circle in the (ρ, μ)-plane centered at $(\rho^0, \mu^0)(0)$ with radius δ; we denote this circle briefly by $C_{\rho;\delta}(0)$.

In this section, for a conveniently chosen φ-modification parameter Φ, the initial value of φ is determined by (34). Fixing $\theta(0) = \frac{\theta_3}{2}$, $\dot{\rho}(0) = 0$ and setting $\rho(0) = \rho^{\epsilon,\delta}(0)$ (cf. (35)), the initial values of $\dot{\varphi}$, $\dot{\theta}$ can be estimated by the method described in Section 4.1. Thus, for any value of $\varphi(0)$ given by (34), infinitely many sets of initial values can be obtained by varying ω on $[0, 2\pi)$ in (35). Experimentally speaking, it suffices for our purposes to vary ω on $\{(\frac{\pi}{2^n})k;\ k = 0, \ldots, (2^{n+1} - 1)\}$, say for $n = 7$. Running our system of ODEs (18) for these sets of initial values and integrating the associated solution shape curves up to the first orthogonally meeting time t_m (cf. (30)), we are interested in recording only the values of $\dot{\rho}$, $\dot{\varphi}$ at $t = t_m$ (cf. Lemma 5.4), and in finding the locus of their corresponding points in the $(\dot{\rho}, \dot{\varphi})$-plane $= \mathbb{R}^2 \subset T_{\bar{\gamma}(t_m)}\tilde{M}$.

We were surprised to find out that the locus of (each set of 2^{7+1}) points $(\dot{\rho}(t_m), \dot{\varphi}(t_m))$ corresponding to each value of $\varphi(0)$, by (34), is *almost* a line in the $(\dot{\rho}, \dot{\varphi})$-plane; to be more precise, the PMCC (Pearson product-moment correlation coefficient) associated to the regression line passing through these points is very close to $+1$ such that we may assume that all these points lie on a line. Let $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ denote the afore-mentioned line in the $(\dot{\rho}, \dot{\varphi})$-plane at $t = t_m$.

Hence, we can define a map transferring data along the solution shape curve γ^* from the δ-circle $C_{\rho;\delta}(0)$ in the (ρ, μ)-plane at $t = 0$ to the line $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ in the $(\dot{\rho}, \dot{\varphi})$-plane at $t = t_m$. We call this map (and its restriction to the discrete subset of the circle $C_{\rho;\delta}(0)$ whose parameter ω varies on $\{(\frac{\pi}{2^n})k;\ k = 0, \ldots, (2^{n+1} - 1)\}$, say for $n = 7$) the *shape transport* on the time-interval $[0, t_m]$, and denote it by $S_{[0,t_m]} : C_{\rho;\delta}(0) \to \ell_{\dot{\rho}\dot{\varphi}}(t_m)$.

Remark 5.6. Replacing the last term at the right-hand side of (33) by another continuous function of ϵ (i.e., changing $\delta =$ radius of the circle $C_{\rho;\delta}(0)$) changes only the distribution of the set of 2^{7+1} points $(\dot{\rho}, \dot{\varphi})$ at $t = t_m$ along the line $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ under the shape transport $S_{[0,t_m]}$, leaving both the $\dot{\varphi}$-intercept and inclination of $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ invariant. But if we modify $\varphi(0)$ appropriately by (34), the line $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ starts moving and sweeps a strip containing the origin of the $(\dot{\rho}, \dot{\varphi})$-plane at $t = t_m$.

On account of the above remark, there exist two values for $\varphi(0)$, by (34), to which correspond two parallel lines $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$ under the shape transport $S_{[0,t_m]}$ whose $\dot{\varphi}$-intercepts have different signs. Therefore, for any given ϵ-dependent mass distribution in Case 1 (resp. Case 2) using the continuity method, there exists a unique value of $\varphi(0)$ less (resp. greater) than $\frac{\pi}{2}$ to which corresponds a line $\ell_{\dot{\rho}\dot{\varphi}}(t_m) \subset (\dot{\rho}, \dot{\varphi})$-plane passing through the origin (cf. Lemma 5.4); i.e., the initial configuration of our m-triangle for any non-uniform mass distribution is an isosceles triangle, and not an Euler configuration. In order to detect the point on the circle $C_{\rho;\delta}(0) \subset (\rho, \mu)$-plane which maps to the origin of the $(\dot{\rho}, \dot{\varphi})$-plane at $t = t_m$ under the shape transport $S_{[0,t_m]}$, and hence to obtain the set of initial values for which the solution shape curve meets the symmetry plane orthogonally at $t = t_m$, it is convenient to focus our attention to a line segment of the desired $\ell_{\dot{\rho}\dot{\varphi}}(t_m) \subset (\dot{\rho}, \dot{\varphi})$-plane around the origin, and to decrease the length of this line segment. To do this, we put some restrictions on the domain of the parameter ω (by increasing n and restricting k to an appropriate subset of $\{0, \ldots, (2^{n+1} - 1)\}$ in $\{(\frac{\pi}{2^n})k;\ k = 0, \ldots, (2^{n+1} - 1)\}$), and change the radius δ of the circle $C_{\rho;\delta}(0)$ appropriately to increase the density of the corresponding points $(\dot{\rho}, \dot{\varphi})$ along the line segment of $\ell_{\dot{\rho}\dot{\varphi}}(t_m)$. In limit, the length of this line segment tends to zero and accordingly $(\dot{\rho}(t_m), \dot{\varphi}(t_m)) \to (0, 0)$ (cf. Fig. 3). This ends the

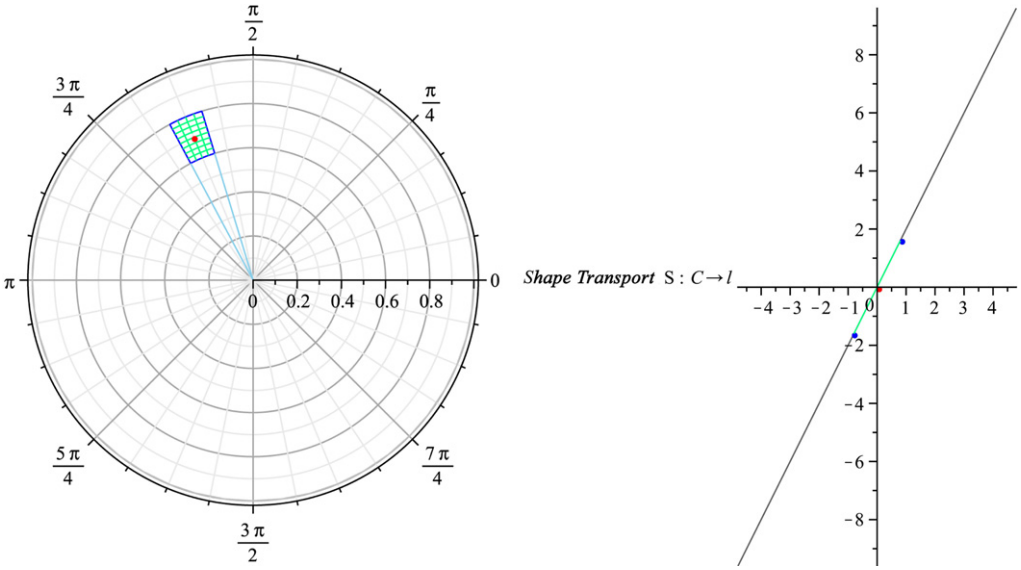

Fig. 3. The shape transport.

procedure of constructing the arc $\gamma^*|_{[0,t_m]}$ of the solution shape curve integrated from $t=0$, starting along the main meridian, up to the first orthogonally meeting time with the canonical symmetry plane at $t=t_m$. It is worth pointing out that both endpoints of the arc $\gamma^*|_{[0,t_m]}$ lie on the upper-(resp. lower-)hemisphere of the 2-sphere M^* in Case 1 (resp. Case 2) of our ϵ-dependent mass distributions; in either case, the arc in fact crosses the equator two times.

Due to the symmetry with respect to the canonical symmetry plane, the next arc of the solution shape curve $\gamma^*|_{[t_m,2t_m]}$ can be constructed which is in fact the reflectional image of $\gamma^*|_{[0,t_m]}$ in the symmetry plane. The important point to note here is that the solution shape curve, after meeting the symmetry plane at $t=t_m$, returns to its starting point at $t=2t_m$. But $\gamma^*|_{[0,t_m]} * \gamma^*|_{[t_m,2t_m]}$ clearly does not give a periodic curve since $\mu(0)\neq\mu(2t_m)$ (cf. Definition 5.3). In order to get a periodic shape curve, the curve must therefore meet the canonical symmetry plane orthogonally at least once more, since otherwise it could not return to its starting point with the same inclination as the initial one.

Lemma 5.7. *The periodic shape curve γ^* associated to a (periodic) three-body trajectory γ meets the canonical symmetry plane orthogonally* exactly *two times on opposite hemispheres of the shape space $M^*=S^2(1)$.*

Proof. For otherwise, if the number of times that the periodic shape curve meets the symmetry plane orthogonally exceeded two, the shape curve would contain at least one loop on $M^*=S^2(1)$ diffeomorphic to a circle in $\mathbb{R}^2$. But eliminating this loop gives another periodic shape curve which is also a solution to the vector ODE

$$\frac{d^2}{dt^2}\gamma^* + \left(2\frac{\dot\rho}{\rho}\right)\frac{d}{dt}\gamma^* - \frac{4}{\rho^3}\nabla U^* = 0$$

with the same initial values $\gamma^*(0)$, $\dot\gamma^*(0)$ as the original one has (cf. [8, Th. 3.9]), contrary to the uniqueness theorem of an ODE. In addition, the meetings occur on opposite hemispheres of the 2-sphere M^*, which is due to Lemmas 5.2 and 5.4. □

Remark 5.8. The above lemma shows that every periodic shape curve consists of exactly two structural arcs, namely restrictions of γ^* to the time-intervals $[t_m, 2t_m+t'_m]$ and $[2t_m+t'_m, 3t_m+2t'_m]$,

where $t = 2t_m + t'_m$ is the second time at which the shape curve meets the symmetry plane orthogonally. Needless to say that (due to the symmetry with respect to the canonical symmetry plane) these two structural arcs are in fact reflectional image of each other in the symmetry plane, and the time it takes we finish traveling along each of which is $t_m + t'_m$; hence the period of the shape curve equals $2(t_m + t'_m)$.

Structural arcs. We continue in the above-mentioned fashion constructing the first structural arc by running our system of ODEs (18) for the terminal data of the arc $\gamma^*|_{[0,t_m]}$ as the initial values, and integrating the associated solution shape curve up to its second time of orthogonally meeting with the canonical symmetry plane.

Let $\Phi_m > 0$ be the φ-modification parameter so that

$$\varphi_{sa}(0) = \varphi(t_m) \pm \Phi_m, \tag{36}$$

where the subscript *sa* refers to the first structural arc for which the origin of time has been translated to the first orthogonally meeting time with the canonical symmetry plane at $t = t_m$.

We shall construct the first structural arc by shooting orthogonally from the first meeting time at $t = t_m$ with the symmetry plane to the second one at $t = 2t_m + t'_m$; i.e., we set $\theta_{sa}(0) = \theta(t_m) = \pi + \frac{\beta_3}{2}$ and $\dot\rho_{sa}(0) = \dot\rho(t_m) = 0 = \dot\varphi(t_m) = \dot\varphi_{sa}(0)$ (cf. Lemma 5.4).

Let us, for sufficiently small $\delta > 0$, set

$$\begin{aligned} \rho_{sa}^{\epsilon,\delta}(0) &= \delta\cos\omega + \rho^\epsilon(t_m),\\ \tilde\mu_{sa}^{\epsilon,\delta}(0) &= \delta\sin\omega, \end{aligned} \tag{37}$$

where ω varies on $[0, 2\pi)$, and $\rho^\epsilon(t_m)$ is the size of our m-triangle at $t = t_m$ for a given ϵ-dependent mass distribution; note that $\mu^\epsilon(t_m)$ is taken to be zero here (cf. (35)).

Setting $\rho_{sa}(0) = \rho_{sa}^{\epsilon,\delta}(0)$, we can calculate $\dot\theta_{sa}(0) = \dot\theta(t_m)$ using the energy integral $h = -0.5 = T - \frac{U^*}{\rho}$ and (23). Thus, for any value of $\varphi_{sa}(0)$ given by (36), infinitely many sets of initial values can be obtained by varying ω on $[0, 2\pi)$ in (37); but, experimentally speaking, it suffices for our purposes to vary ω on $\{(\frac{\pi}{2^n})k;\ k = 0, \dots, (2^{n+1} - 1)\}$, say for $n = 7$.

In much the same way as used in constructing the arc $\gamma^*|_{[0,t_m]}$ of the shape curve up to the first orthogonally meeting time at $t = t_m$, here we can also make beneficial use of the shape transport to construct the first structural arc.

Let $C_{\rho;\delta}(t_m)$ denote both the δ-circle $\{(\rho_{sa}^{\epsilon,\delta}, \tilde\mu_{sa}^{\epsilon,\delta})(0);\ \omega \in [0, 2\pi)\}$ in the (ρ, μ)-plane $= \mathbb{R}^2 \subset T_{\bar\gamma(t_m)}\bar M$ centered at $(\rho^\epsilon(t_m), 0)$, cf. (37), as well as its discrete subset with ω varying on $\{(\frac{\pi}{2^7})k;\ k = 0, \dots, (2^{7+1} - 1)\}$.

Running our system of ODEs (18) for the above-mentioned sets of initial values and integrating the associated solution shape curve along the first structural arc on $[t_m, 2t_m + t'_m]$, we were surprised at finding out that to each set of 2^{7+1} points

$$\big(\dot\rho_{sa}(t_m + t'_m), \dot\varphi_{sa}(t_m + t'_m)\big) = \big(\dot\rho(2t_m + t'_m), \dot\varphi(2t_m + t'_m)\big)$$

there corresponds a regression line with the PMCC very close to -1 such that we may assume that all these points lie on a line, denoted by $\ell_{\dot\rho\dot\varphi}(2t_m + t'_m)$, in the $(\dot\rho, \dot\varphi)$-plane $= \mathbb{R}^2 \subset T_{\bar\gamma(2t_m + t'_m)}\bar M$.

Applying the same arguments as Remark 5.6, with the time-interval $[0, t_m]$ replaced by $[t_m, 2t_m + t'_m]$, we can conclude by the continuity method that there exists a unique value of $\varphi_{sa}(0) = \varphi(t_m) \pm \Phi_m$ to which corresponds the line $\ell_{\dot\rho\dot\varphi}(2t_m + t'_m) \subset (\dot\rho, \dot\varphi)$-plane passing through the origin (cf. Lemma 5.4). It follows that the point on the circle $C_{\rho;\delta}(t_m) \subset (\rho, \mu)$-plane which maps to the origin of the $(\dot\rho, \dot\varphi)$-plane at $t = 2t_m + t'_m$ under the shape transport

$$S_{[t_m, 2t_m + t'_m]} : C_{\rho;\delta}(t_m) \to \ell_{\dot\rho\dot\varphi}\big(2t_m + t'_m\big),$$

and hence the set of initial values for which the first structural arc of the shape curve starts orthogonally to the symmetry plane at $t = t_m$ and meets the symmetry plane again orthogonally on the opposite hemisphere at $t = 2t_m + t'_m$, can be detected by the same method as in constructing the arc $\gamma^*|_{[0,t_m]}$ of the shape curve.

This finishes procedure of constructing the first structural arc $\gamma^*|_{[t_m,2t_m+t'_m]}$ of the shape curve which crosses the equator at one and only one eclipse point, other than the Euler point e_3, for any non-uniform mass distribution with $m_1 = m_2$. Consequently, we have constructed the periodic shape curve $\gamma^*(t)$ associated to the three-body trajectory $\gamma(t)$ since the second structural arc $\gamma^*|_{[2t_m+t'_m,3t_m+2t'_m]}$ is the reflectional image of the first one in the canonical symmetry plane, which is due to the symmetry with respect to the symmetry plane.

Evidently, for any periodic three-body motion, the associated shape curve and the moduli curve are both periodic; but the converse may not hold. However, Lemma 5.1 shows that the time parametrization of the shape curve $\gamma^*(t)$ is dictated by the geometry of the oriented geometric (i.e., arc-length parametrized) shape curve γ^*; hence periodicity of the shape curve is equivalent to that of the moduli curve. Therefore, for any given ϵ-dependent mass distribution, we have succeeded in constructing a periodic moduli curve (i.e., a *relative periodic* orbit) which has the reflectional symmetry with respect to the canonical symmetry plane.

5.3. Periodicity of the three-body trajectories

What is left is to show that the relative periodic three-body trajectory $\gamma(t)$, for any given ϵ-dependent mass distribution, is actually periodic. To this end, let $D \subset M^* = S^2(1)$ be the region enclosed by the associated periodic shape curve whose boundary is oriented in accordance with the oriented geometric shape curve γ^*. Clearly, any periodic shape curve is closed; thus to the endpoints of the three-body trajectory $\gamma(t)$ associated to one period of the shape curve there correspond congruent m-triangles which therefore differ only by a (total) rotation angle $\Delta\psi$ to be calculated as a line integral along the shape curve (i.e., along the boundary ∂D), on account of the kinematic Gauss–Bonnet theorem (cf. [9, Th. C2]). Since the above-mentioned line integral (applying Green's theorem) can be expressed as a surface integral which yields $\Delta\psi$ equals half the *signed area* of the region D (indeed, $\Delta\psi = \frac{1}{2r^2}\,Area(D)$ for $D \subset M^* = S^2(r)$), it follows that the motion $\gamma(t)$ is periodic if and only if the total rotation angle $\Delta\psi = \frac{1}{2}\,Area(D)$ is a rational multiple of 2π; in particular, $\Delta\psi = 0$ implies $\gamma(t)$ is periodic.

We shall now prove that the signed area of the region D enclosed by (or, the total rotation angle $\Delta\psi$ associated to) our periodic shape curve for any given ϵ-dependent mass distribution is zero. To this end, we project the region $D \subset M^* = S^2(1)$ to the open disk of radius 2 centered at the origin of the plane $\mathbb{R}^2$ tangent to the 2-sphere M^* at its south pole $\mathbf{S}$ under the *Lambert azimuthal area-preserving projection map*, and rotate the projected image by angle $\frac{-\beta_3}{2}$ in the Euclidean plane. In the spherical polar coordinates (φ, θ) on the 2-sphere M^* (with φ as the zenith, θ the azimuth and $\mathbf{N}$ the north pole) and the standard polar coordinates (r, θ) on the open disk, the Lambert projection map reads

$$\begin{aligned} L : S^2(1) \setminus \mathbf{N} &\to B(o, 2), \\ (\varphi, \theta) &\mapsto \left(2\cos\left(\frac{\varphi}{2}\right), \theta\right) \end{aligned} \tag{38}$$

which is actually a diffeomorphism. Computing the area element of the 2-sphere M^* when parametrized by the inverse of the projection L, we can see that the Lambert azimuthal projection map is an area-preserving one.

By (38) it is evident that the Lambert projection map L projects each point on the 2-sphere (except the north pole) to the open disk along a circular arc centered at the south pole. In particular, it sends the south pole $\mathbf{S}$ to the origin of $\mathbb{R}^2$, and projects the equator to the circle of radius $\sqrt{2}$ centered at the origin which is called the *Lambert circle* (cf. Fig. 4).

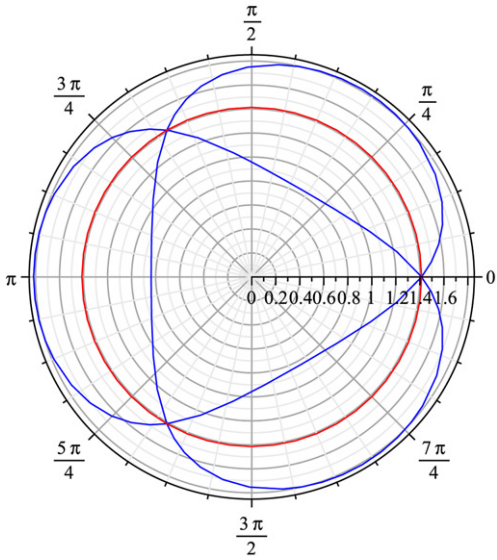

Fig. 4. The Lambert projection. The projected shape curve is in blue and the Lambert circle in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

It follows that the main meridian is projected, up to a rotation by angle $\frac{-\beta_3}{2}$, to the diameter of the circle $\partial B(o,2) \subset \mathbb{R}^2$ laying on the x-axis of the Cartesian plane $\mathbb{R}^2$. Furthermore, by (38) and due to the symmetry of our periodic solution shape curve $\gamma^* \subset M^*$ (for any given mass distribution) with respect to the canonical symmetry plane containing the main meridian, the region $D \subset M^* = S^2(1)$ enclosed by γ^* is projected under $(R_{\frac{-\beta_3}{2}} \circ L)$ to a region in $\mathbb{R}^2$ enclosed by the closed curve $L(\gamma^*) \subset B(o,2) \subset \mathbb{R}^2$ which is symmetric with respect to the x-axis, where $R_{\frac{-\beta_3}{2}} \circ L$ is the composition of the Lambert projection map and the rotation by angle $\frac{-\beta_3}{2}$.

Consequently, the signed area of the region $(R_{\frac{-\beta_3}{2}} \circ L)(D) \subset B(o,2) \subset \mathbb{R}^2$ is zero and so is that of $D \subset M^* = S^2(1)$ (i.e., $Area(D) = 0$) since $R_{\frac{-\beta_3}{2}} \circ L$ is an area-preserving bijection. This finishes the proof of periodicity of our three-body trajectories $\gamma(t)$ for any given mass distribution with $m_1 = m_2$. In the case of non-uniform mass distributions with $m_1 = m_2$, the bodies move on three different deformed 8-shaped periodic orbits of the same period as that of the solution curve $\gamma(t)$ in the inertial plane; moreover, γ has the reflectional symmetry with respect to the symmetry axis of the third body's periodic orbit.

5.4. Individual torques: Star-convexity of the 8-shaped orbits' lobes

First recall that *torque* (also called *moment of force*) is, by definition, the tendency of a force to rotate an object about an axis. It can be defined as the cross product of the particle's position vector (in a reference frame) and the force acting on the particle (cf. [7,12]).

In the study of m-triangles whose geometry depends on the mass distributions, it is natural to apply the following area law

$$\textit{Ceva-area law}: \quad \Delta_j = m_j \Delta, \tag{39}$$

where Δ is the area of the m-triangle with vertexes to which the point masses m_i's are attached, and Δ_j is that of the m-triangle with jth point mass replaced by the center of mass (i.e., by the origin in our barycentric coordinate system).

As a simple application of the Ceva-area law we will use the following lemma concerning the individual torque of each body m_i due to the resultant Newtonian gravitational force acting on the point mass m_i:

Lemma 5.9. *Let $\boldsymbol{\tau}_i$ be the individual torque of the point mass m_i, $i = 1, 2, 3$, with respect to the center of mass. Then*

$$\boldsymbol{\tau}_i = \dot{\boldsymbol{\Omega}}_i = 2m_1m_2m_3\Delta\left(\frac{1}{r_{i,i+1}^3} - \frac{1}{r_{i,i+2}^3}\right)\mathbf{n}, \quad \textit{indexes} \pmod 3,$$

where $\mathbf{n}$ is the unit normal vector so that $(\mathbf{a}_1, \mathbf{a}_2, \mathbf{n})$ is a right-handed frame, and $\boldsymbol{\Omega}_i = m_i(\mathbf{a}_i \times \dot{\mathbf{a}}_i)$ is the individual angular momentum of m_i.

Proof. Let $\mathbf{F}_{12}$ and $\mathbf{F}_{13}$ be the gravitational forces respectively due to the point masses m_2 and m_3 acting on m_1, namely

$$\mathbf{F}_{1j} = \frac{m_1m_j}{r_{1j}^3}(\mathbf{a}_j - \mathbf{a}_1), \quad j = 2, 3$$

which read the Newton's equations of motion (2): $m_1\ddot{\mathbf{a}}_1 = \mathbf{F}_{12} + \mathbf{F}_{13}$. Then, by definition of torque and the Ceva-area law (39),

$$\begin{aligned}\boldsymbol{\tau}_1 = \dot{\boldsymbol{\Omega}}_1 &= \mathbf{a}_1 \times (m_1\ddot{\mathbf{a}}_1) = \mathbf{a}_1 \times (\mathbf{F}_{12} \times \mathbf{F}_{13}) = \frac{m_1m_2}{r_{12}^3}(\mathbf{a}_1 \times \mathbf{a}_2) + \frac{m_1m_3}{r_{13}^3}(\mathbf{a}_1 \times \mathbf{a}_3)\\ &= \left(\frac{m_1m_2}{r_{12}^3}2m_3\Delta\right)\mathbf{n} - \left(\frac{m_1m_3}{r_{13}^3}2m_2\Delta\right)\mathbf{n} = 2m_1m_2m_3\Delta\left(\frac{1}{r_{12}^3} - \frac{1}{r_{13}^3}\right)\mathbf{n};\end{aligned}$$

and the other individual torques can be obtained similarly to above. □

Remark 5.10. Geometrically speaking, the individual torque $\boldsymbol{\tau}_i = 0$ (at time $t = t_0$) if and only if

(i) $r_{ij} = r_{ik}$ (i.e., the m-triangle at $t = t_0$ has an isosceles configuration with the ith mass on the symmetry axis of the triangle), or
(ii) $\Delta = 0$ (i.e., the m-triangle at $t = t_0$ is a degenerate one: the three masses aligned).

A closed curve in the Euclidean plane is called *star-convex* with respect to a point $P_\star$ inside it, or on its boundary, if every line segment starting from $P_\star$ intersects the curve at most once. If a closed curve is smooth (except at its endpoints), this amounts to the assertion that – when written in the polar coordinates $(r(t), \theta(t))$ – the function $\theta(t)$ is strictly monotone and does not vary by more than 2π. Since in the polar coordinates the individual angular momenta can be expressed as $\boldsymbol{\Omega}_i = m_i(r_i^2\dot{\theta}_i)\mathbf{n}$, the star-convexity of each lobe of our deformed 8-shaped orbit $\boldsymbol{b}_i$ is thus equivalent to $\boldsymbol{\Omega}_i \neq 0$ on that lobe. In what follows, the bivector $\boldsymbol{\Omega}_i$ or $\boldsymbol{\tau}_i$ is called positive if it is a positive multiple of the vector $\mathbf{n}$ (cf. Lemma 5.9).

On account of Lemma 5.9 and the above remark, we are interested in finding isosceles configurations and places where the three masses are aligned (i.e., we arrive at a degenerate m-triangle) as we travel our periodic solution curve $\gamma = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$, where $\boldsymbol{b}_i$ is the periodic orbit (of the same period as that of γ) along which the mass m_i moves. Note that as we travel our solution curve in the inertial plane the following hold:

(i) When we arrive at an *isosceles configuration of type i* (i.e., an isosceles m-triangle with the mass m_i on its symmetry axis), the associated shape curve γ^* and the modulo curve $\bar{\gamma}$ meet the

plane in the modulo space $\bar{M}$ which is perpendicular to the equatorial plane and passes through the binary collision point b_{jk}, $\{i, j, k\} = \{1, 2, 3\}$ (for $i = 3$, this plane is the canonical symmetry plane, cf. Definition 4.1). The isosceles configurations of type i, with the exception of the initial configuration for non-uniform mass distributions, will be denoted by Isc_i^u and Isc_i^l depending on whether the associated shape point is on the *u*pper or *l*ower hemisphere of the 2-sphere M^*; the initial configuration for any non-uniform mass distribution is in fact an isosceles one of type 3 and will be denoted by Isc_3^o whose associated shape point is on the upper-(resp. lower-)hemisphere of M^* in Case 1 (resp. Case 2) of our mass distributions.

(ii) If we arrive at a degenerate m-triangle, the associated shape curve γ^* meets the equator of the 2-sphere M^* at an *eclipse point*, briefly denoted by *Ec*, which will be labeled with an alphabetic index appropriately.

The fact that the signed area enclosed by (and hence the total rotation angle associated to) our periodic shape curve equals zero implies that the congruent m-triangles corresponded to the endpoints of the periodic solution curve $\gamma(t)$ associated to one period of the shape curve actually coincide.

Given any non-uniform mass distribution in Case 1, as we travel the first structural arc of our modulo curve $\bar{\gamma}$ (i.e., from Isc_3^u to Isc_3^l), the three masses (each of which moves along its own periodic orbit in the inertial plane) pass through isosceles configurations and aligned in the following order:

$$Isc_3^u, Ec_a, Isc_2^l, Ec_b, Isc_3^o, Isc_1^u, Ec_c, Isc_3^l; \tag{IscEc1}$$

and as we travel the second structural arc of $\bar{\gamma}$ (i.e., from Isc_3^l to Isc_3^u) to complete its first period up to a time translation, the masses pass through

$$Isc_3^l, Ec_d, Isc_2^u, Isc_3^o, Ec_e, Isc_1^l, Ec_f, Isc_3^u. \tag{IscEc2}$$

Hence, the orbit $\boldsymbol{b}_3$ has a self-intersection point O_3 to which there corresponds the isosceles configuration Isc_3^o. Moreover, choosing the origin of time $t = 0$ to correspond temporarily to being in an isosceles configuration Isc_i^u (or Isc_i^l), $i = 1, 2$, we can assert that the orbit $\boldsymbol{b}_i$ has also a self-intersection point at O_i to which there correspond the isosceles configurations Isc_i^u and Isc_i^l, $i = 1, 2$ (cf. Fig. 5).

Likewise for any given non-uniform mass distribution in Case 2, interchanging the superscript u and l in (*IscEc*1) and (*IscEc*2), we have the list of succeeding configurations formed by the masses in the inertial plane as we travel the structural arcs of our modulo curve $\bar{\gamma}$. Hence, we can deduce that the periodic orbit $\boldsymbol{b}_i$ traversed by m_i in the inertial plane has a self-intersection point O_i.

Corollary 5.11. *Each lobe of the self-intersecting periodic orbit $\boldsymbol{b}_i$, $i = 1, 2, 3$, is star-convex.*

Proof. Translating all the $\boldsymbol{b}_i$'s in a fixed direction keeps both the configurations of m-triangles and the integral of vanishing (total) angular momentum invariant, the latter is due to the fact that $\sum_1^3 m_i \dot{\mathbf{a}}_i = 0$. Hence, there is no loss of generality in assuming the origin to be any arbitrary point in the plane. Here, by *the first* (resp. *the second*) *lobe* of $\boldsymbol{b}_i$ we mean the lobe on which m_i passes through the isosceles configuration Isc_3^u (resp. Isc_3^l).

Let (m_1, m_2, m_3) be a given ϵ-dependent mass distribution as in Case 1 with $0 < \epsilon < \frac{1}{3}$, and $\gamma = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$ be our periodic solution curve in the inertial plane.

The proof falls naturally into several parts.

The first lobe of $\boldsymbol{b}_1$. Let the origin be the intersection point, denoted by $P_{\star 1}$, of the tangent line to $O_1 = Isc_1^l$ and the ray from the center of mass which meets $\boldsymbol{b}_1$ at Ec_f (cf. (*IscEc*2)). Let q_1 be the (first) point at which the tangent line to O_1 meets the boundary of the first lobe of $\boldsymbol{b}_1$, and $\boldsymbol{\tau}_{t1}$ and $\boldsymbol{\Omega}_{t1}$ be respectively the individual torque and angular momentum of m_1 with respect to the (translated) origin $P_{\star 1}$. It is immediate that

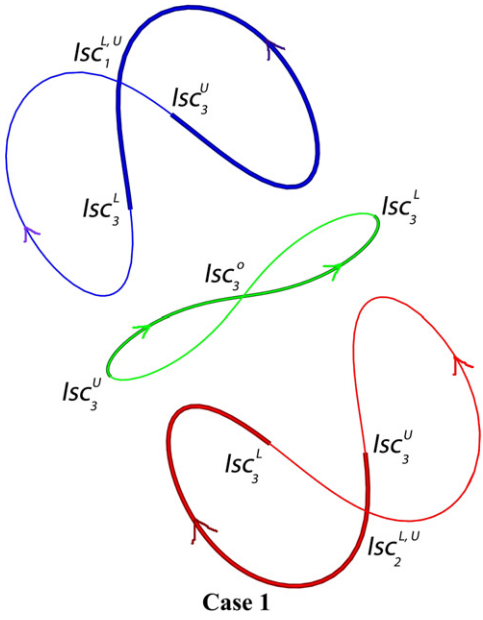

Fig. 5. Consecutive configurations. As we travel the first structural arc of our modulo curve, each of the three masses traverses the bold arc along its own periodic orbit. For convenience, the orbits of m_1 (in blue, at the top), m_2 (in red, at the bottom), and that of m_3 (in green, in the middle) have been plotted in three different (translated) Cartesian coordinate systems. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\boldsymbol{\tau}_{t1} = (\mathbf{a}_1 - P_{\star 1}) \times (m_1 \ddot{\mathbf{a}}_1) = \boldsymbol{\tau}_1 - P_{\star 1} \times (m_1 \ddot{\mathbf{a}}_1),$$
$$\boldsymbol{\Omega}_{t1} = (\mathbf{a}_1 - P_{\star 1}) \times (m_1 \dot{\mathbf{a}}_1) = \boldsymbol{\Omega}_1 - P_{\star 1} \times (m_1 \dot{\mathbf{a}}_1).$$

On account of the Newton's equations of motion (2), the acceleration $\ddot{\mathbf{a}}_1$ is a linear combination of the lateral vectors $(\mathbf{a}_2 - \mathbf{a}_1)$ and $(\mathbf{a}_3 - \mathbf{a}_1)$ of our m-triangle with positive coefficients. As m_1 travels the curve joining $O_1 = Isc_1^l$ to Ec_f (excluding the endpoint) on $\boldsymbol{b}_1$, boundary of the first lobe of $\boldsymbol{b}_2$ from Isc_1^l to Ec_f and that of $\boldsymbol{b}_3$ are respectively swept by m_2 and m_3; it follows that the angle between $(\mathbf{a}_1 - P_{\star 1})$ and $(m_1 \ddot{\mathbf{a}}_1)$, denoted by α_1, does not exceed π (i.e., it is always an acute, right or an obtuse angle), and hence $\boldsymbol{\tau}_{t1}$ is positive on this curve. Notice that α_1 is particularly zero at Ec_f, then so is $\boldsymbol{\tau}_{t1}$.

Throughout the proof, the angle α_i, $i = 1, 2$, between two vectors $(\mathbf{a}_i - P_{\star i})$ and $(m_i \ddot{\mathbf{a}}_i)$ is measured as usual by the amount of rotation about $P_{\star i}$ in the positive trigonometric direction (i.e., counterclockwise) which is required to bring $(\mathbf{a}_i - P_{\star i})$ into correspondence with $(m_i \ddot{\mathbf{a}}_i)$.

When m_1 continues to move along the first lobe through the point q_1, α_1 is again an acute, right or obtuse angle which vanishes at q_1; but α_1 is a reflex angle when m_1 is moving from q_1 to $O_1 = Isc_1^u$.

Therefore, as m_1 travels the first lobe of $\boldsymbol{b}_1$, its individual torque $\boldsymbol{\tau}_{t1}$ stays non-negative between $O_1 = Isc_1^l$ and q_1, vanishes only at Ec_f and q_1, and is negative between q_1 and $O_1 = Isc_1^u$.

It follows that the individual angular momentum $\boldsymbol{\Omega}_{t1}$ increases as m_1 travels from $O_1 = Isc_1^l$ to q_1, and decreases as m_1 moves from q_1 to $O_1 = Isc_1^u$. To conclude that $\boldsymbol{\Omega}_{t1}$ stays strictly positive on the first lobe of $\boldsymbol{b}_1$ (excluding the starting point Isc_1^l), it remains to notice that $\boldsymbol{\Omega}_{t1} = 0$ at Isc_1^l since $(\mathbf{a}_1 - P_{\star 1})$ and $(m_1 \dot{\mathbf{a}}_1)$ are collinear there, and that $\boldsymbol{\Omega}_{t1}$ is positive at Isc_1^u since at that point cannot the angle between the tangent vectors $(\mathbf{a}_1 - P_{\star 1})$ and $(m_1 \dot{\mathbf{a}}_1)$ exceed π. Consequently, the first lobe of $\boldsymbol{b}_1$ is star-convex.

The second lobe of $\boldsymbol{b}_1$. The star-convexity of this lobe can be concluded in much the same way as above, the only difference being in the choice of the (translated) origin $P_{\star 1}$. We now choose the origin $P_{\star 1}$

to be the intersection point of the tangent line to $O_1 = Isc_1^u$ and the ray from the center of mass which meets $\boldsymbol{b}_1$ at Ec_c (cf. (*IscEc*1)), and again let q_1 be the (first) point at which the tangent line to $O_1 = Isc_1^u$ meets the boundary of the second lobe of $\boldsymbol{b}_1$. It follows that the individual angular momentum $\boldsymbol{\Omega}_{t1}$ decreases as m_1 travels from $O_1 = Isc_1^u$ to q_1 – passing through Ec_c at which $\boldsymbol{\tau}_{t1}$ vanishes – and increases as m_1 moves from q_1 to $O_1 = Isc_1^l$; hence the star-convexity of this lobe is a consequence of the facts that $\boldsymbol{\Omega}_{t1} = 0$ at Isc_1^u since $(\mathbf{a}_1 - P_{\star 1})$ and $(m_1\dot{\mathbf{a}}_1)$ are collinear there, and that $\boldsymbol{\Omega}_{t1}$ is negative at Isc_1^l since at that point will the angle between the tangent vectors $(\mathbf{a}_1 - P_{\star 1})$ and $(m_1\dot{\mathbf{a}}_1)$ be a reflex one.

The first lobe of $\boldsymbol{b}_3$. We now choose the origin to be the self-intersection point O_3 of the periodic orbit $\boldsymbol{b}_3$ (i.e., we set $P_{\star 3} = O_3$). To measure the angle α_3 between $(\mathbf{a}_3 - P_{\star 3})$ and $(m_3\ddot{\mathbf{a}}_3)$, we adhere to the same standard convention as in the proof of star-convexity of the first lobe of $\boldsymbol{b}_1$.

As m_3 travels the boundary of this lobe from $O_3 = Isc_3^o$ to Isc_3^u via Isc_1^l (cf. (*IscEc*2)) which corresponds to the curve joining Isc_3^o to Isc_3^u along the second structural arc of the modulo curve $\bar{\gamma}$ in the modulo space, m_1 starts moving from Isc_3^o on the boundary of the second lobe of $\boldsymbol{b}_1$ and arrives at Isc_3^u on the first lobe via the self-intersection point $O_1 = Isc_1^l$, and the boundary of the first lobe of $\boldsymbol{b}_2$ is swept by m_2 from Isc_3^o to Isc_3^u. Hence α_3 is always a reflex angle as m_3 is on the curve between $O_3 = Isc_3^o$ and Isc_3^u which contains the point Isc_1^l, and thus $\boldsymbol{\tau}_{t3}$ stays non-positive on this curve; in fact, $\boldsymbol{\tau}_{t3}$ vanishes only at $O_3 = Isc_3^o$ and Isc_3^u.

Therefore, $\boldsymbol{\Omega}_{t3}$ decreases from the value 0 at O_3 as m_3 travels the boundary of this lobe from $O_3 = Isc_3^o$ to Isc_3^u via Isc_1^l, and hence increases to the value 0 at O_3 as m_3 does so from Isc_3^u to O_3 via Isc_2^l (cf. (*IscEc*1)), which the latter assertion is due to the fact that these two paths are reflectional image of each other with respect to the symmetry axis of $\boldsymbol{b}_3$. It follows that $\boldsymbol{\Omega}_{t3}$ stays strictly negative on the first lobe (excluding the point O_3). Consequently, the first lobe of $\boldsymbol{b}_3$ is star-convex.

The second lobe of $\boldsymbol{b}_3$. Analysis similar to that in the proof of star-convexity of the first lobe of $\boldsymbol{b}_3$ shows that $\boldsymbol{\Omega}_{t3}$ increases from the value 0 at O_3 as m_3 travels the boundary of the second lobe from $O_3 = Isc_3^o$ to Isc_3^l via Isc_1^u, and decreases to the value 0 at O_3 as m_3 does so from Isc_3^l to O_3 via Isc_2^u. It follows that it stays strictly positive on the second lobe (excluding the point O_3). Consequently, the second lobe of $\boldsymbol{b}_3$ is also star-convex.

The lobes of $\boldsymbol{b}_2$. The same conclusion can be drawn for these lobes, which is due to the fact that the 8-shaped periodic orbits $\boldsymbol{b}_1$, $\boldsymbol{b}_2$ are reflectional image of each other with respect to the symmetry axis of the third one $\boldsymbol{b}_3$, and that the lobes of $\boldsymbol{b}_1$ are star-convex which has already been proved above.

Similar arguments apply to our periodic solution curve $\gamma = (\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3)$ of the three-body problem with any given ϵ-dependent mass distribution as in Case 2 with $0 < \epsilon < \frac{1}{6}$, the only difference being in the signum analysis of the individual torques and angular momenta of the three bodies.

This finishes our proof of the corollary. □

Conjecture 5.12. *Based on some numerical experiments, it is conjectured that the lobes of* $\boldsymbol{b}_3$*, for any given non-uniform mass distribution* (m_1, m_2, m_3) *as in the Theorem* (*see Section* 3)*, are in fact convex, while those of* $\boldsymbol{b}_1$ *and* $\boldsymbol{b}_2$ *are not.*

The structure of our family of periodic orbits. As a matter of fact, by modification of the masses using the method of analytic continuation, one can construct a family of periodic orbits if the quest starts with an already known periodic orbit.

It is worth pointing out that the structure of our family of periodic orbits is in conformity with the *principle of natural termination*, obtained empirically by Strömgren [16,17] on the basis of numerical explorations and proved by Wintner [20] and Birkhoff [2]. The essence of the principle is as follows: Starting from any given orbit in a family (in our study, the figure-8 orbit E), one can move along the family in two and only two directions in each of which the family has a *natural termination*; by a natural termination in a family we mean a termination without meeting an "old member" of the family.

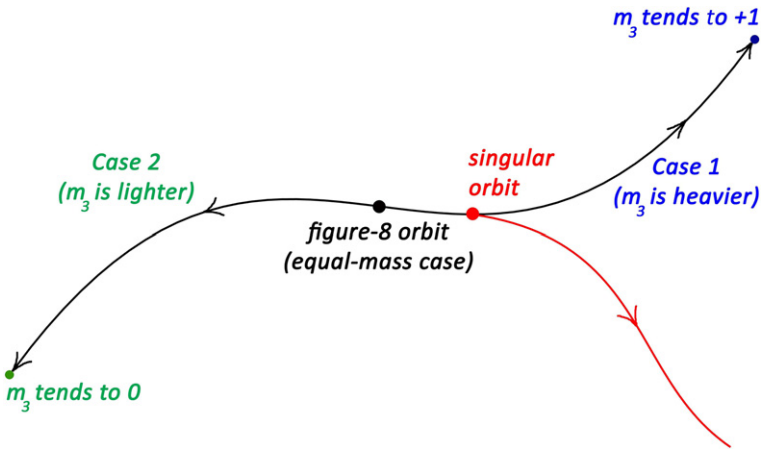

Fig. 6. The principle of natural termination.

In addition, we do not encounter any singularity when we are moving along our family of periodic orbits. For otherwise, if there is a singular orbit then – by principles of the method of analytic continuation – there will be a branching in the family at the point corresponding to the singular orbit (cf. Fig. 6); contrary to the essence of the principle of natural termination, because there will be more than two directions to move along the family starting from the singular orbit.

Note that we have actually proved that moving along the line segment $m_1 = m_2$ on the mass space $\{(m_1, m_2, m_3) \in \mathbb{R}^3 \mid m_i > 0,\ m_1 + m_2 + m_3 = 1\}$ in the two directions with respect to (the heavier or lighter) m_3, yields two continuous families of periodic three-body motions with non-uniform mass distributions of the form γ as in part (A) of the Theorem which connect at the celebrated figure-8 orbit E in the equal-mass case, and yield a continuous family of vanishing angular momentum periodic solutions of the three-body problem in the plane.

To treat the limiting cases of the constructed family of periodic three-body motions we make the following definition.

The shape curve γ^* associated to a solution curve γ of the three-body problem is called *exceptional* if it is confined to a geodesic circle, or it consists of only a single point (which is necessarily one of the five critical points of the shape potential U^*, namely the two minima called Lagrange points and the three saddle points called Euler points) on the 2-sphere M^* (cf. [8, Sec. 3.4]).

At one end of the family, when the two equal masses are infinitesimal and the third one reaches the value of $+1$, the equal masses describe an elliptic Keplerian motion on a circle centered at the origin, but in opposite directions around m_3 which rests at the origin (i.e., we have a double Kepler problem). In fact, the infinitesimal equal masses move on the two different semicircles symmetric with respect to the y-axis; the three-body motion is *shape invariant* in the sense that the three masses always form isosceles configurations of type 3.

At the other end of the family, when the third mass is infinitesimal, orbits of the bodies are flattened out so that the equal masses move along the x-axis and m_3 moves along the y-axis (i.e., we have a special case of periodic solutions of a restricted 3-body problem). Moreover, the motion is shape invariant: the masses form an isosceles configuration of type 3 at every instant of time. In fact, in either limiting case, the associated shape curve is exceptional and confined to the main meridian passing through the binary collision point b_{12} (and the Euler point e_3) on the shape space M^*.

Acknowledgments

The author is greatly indebted to Prof. Wu-Yi Hsiang for suggesting the problem and for many stimulating conversations. My deep gratitude and appreciation go to the Department of Mathematics

at the University of California at Berkeley (UCB), where the main parts of this paper were written, for the invitation and gracious hospitality during my research stay at UCB; warm thanks to my advisor Eldar Straume for many enlightening discussions, and to my home university NTNU for the financial support.

Appendix A. The mass space

The mass space $\{(m_1, m_2, m_3) \in \mathbb{R}^3 \mid m_i \geqslant 0,\ m_1 + m_2 + m_3 = 1\}$ can be given by an equilateral triangle with side length equal to $2/\sqrt{3}$ (cf. Fig. A.1). That is to say, to a given point on the mass space there corresponds a mass distribution (with non-negative masses) (m_1, m_2, m_3) which is normalized so that $m_1 + m_2 + m_3 = 1$. Labeling the vertexes of this equilateral triangle with 1, 2, 3, we can assert that m_i equals the height of the given point on the mass triangle from the opposite side of the vertex i; hence m_i takes the value 0 on the opposite side of the vertex i, and does its maximum value $+1$ at the vertex i.

Moving along the line segment $m_1 = m_2$ (which is perpendicular to the side on which m_3 is zero) in the mass space and using the method of analytic continuation (MAC) starting from the celebrated figure-8 orbit E, we have exhibited a new family of periodic orbits for the Newtonian three-body problem one of whose members is the orbit E; hence this family may be referred to as the *E-family of periodic orbits*.

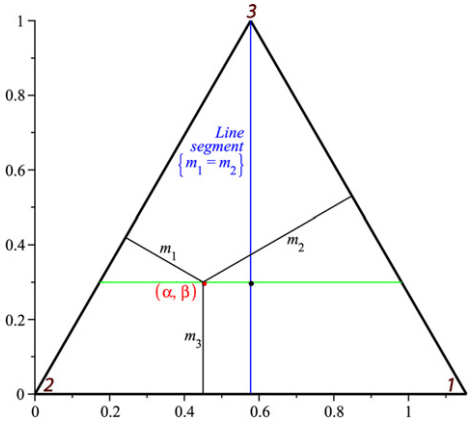

Fig. A.1. The mass space.

One can continue in this fashion obtaining even more interesting families of periodic orbits: fixing $\bar{\epsilon} \in [\frac{-1}{6}, \frac{1}{3}]$ (whose absolute value $|\bar{\epsilon}|$ gives our *mass parameter* ϵ), and using MAC starting from the member of our E-family with $\bar{\epsilon}$-dependent mass distribution

$$m_1 = m_2 = \frac{1}{3} - \bar{\epsilon}, \qquad m_3 = \frac{1}{3} + 2\bar{\epsilon},$$

we can start looking for a new family of periodic orbits by moving along the line segment which passes through the point on the mass triangle corresponding to the above $\bar{\epsilon}$-dependent mass distribution $(\frac{1}{3} - \bar{\epsilon}, \frac{1}{3} - \bar{\epsilon}, \frac{1}{3} + 2\bar{\epsilon})$, and is parallel to the side of the mass triangle on which m_3 is zero. To every point (α, β) on this line segment there corresponds the following mass distribution

$$m_1 = \frac{\sqrt{3}\alpha - \beta}{2}, \qquad m_2 = 1 - \frac{\sqrt{3}\alpha + \beta}{2}, \qquad m_3 = \beta.$$

On the other hand, for the fixed $\bar{\epsilon}$ and any given $|\delta| \leqslant (\frac{1}{3} - \bar{\epsilon})$, to the mass distribution defined by

$$m_1 = \left(\frac{1}{3} - \bar{\epsilon}\right) + \delta, \qquad m_2 = \left(\frac{1}{3} - \bar{\epsilon}\right) - \delta, \qquad m_3 = \frac{1}{3} + 2\bar{\epsilon},$$

there corresponds the point $(\alpha, \beta) = (\frac{1+2\delta}{\sqrt{3}}, \frac{1}{3} + 2\bar{\epsilon})$ on this line segment which is parallel to the opposite side of the vertex 3 of our mass triangle.

References

[1] A. Back, W.Y. Hsiang, Equivariant geometry and Kervaire spheres, Trans. Amer. Math. Soc. 304 (1987) 207–227.
[2] G.D. Birkhoff, Sur le problème restreint des trois corps (second mémoire), in: Collect. Math. Papers, vol. 2, Dover Publications, New York, 1968, pp. 668–709.
[3] A. Chenciner, R. Montgomery, A remarkable periodic solution of the three-body problem in the case of equal masses, Ann. of Math. 152 (2000) 881–901.
[4] A. Deprit, J. Henrard, Construction of orbits asymptotic to a periodic orbit, Astronom. J. 74 (1969) 308–316.
[5] F. Diacu, P. Holmes, Celestial Encounters: The Origins of Chaos and Stability, Princeton University Press, 1996.
[6] T. Fujiwara, H. Fukuda, A. Kameyama, H. Ozaki, M. Yamada, Synchronized similar triangles for three-body orbits with zero angular momentum, J. Phys. A 37 (2004) 10571–10584.
[7] D. Halliday, R. Resnick, Fundamentals of Physics, John Wiley and Sons, Inc., 1970, pp. 184, 185.
[8] W.Y. Hsiang, E. Straume, Global geometry of 3-body motions with vanishing angular momentum I, Chin. Ann. Math. Ser. B 29 (2008) 1–54.
[9] W.Y. Hsiang, E. Straume, Kinematic geometry of triangles and the study of the three-body problem, Lobachevskii J. Math. 25 (2007) 9–130.
[10] J.L. Lagrange, Essai sur le problème de trois corps, Ouvres 6 (1772) 229–324.
[11] H. Poincaré, Les méthodes nouvelles de la Mécanique Céleste, Dover Publications, New York, 1957, pp. 1–174.
[12] R.A. Serway, J.W. Jewett Jr., Physics for Scientists and Engineers, sixth ed., Brooks Cole, 2003.
[13] C. Simó, Dynamical properties of the figure eight solution of the three-body problem, in: Contemp. Math., vol. 292, Amer. Math. Soc., 2002, pp. 1–20.
[14] E. Straume, A geometric study of many body systems, Lobachevskii J. Math. 24 (2006) 73–134.
[15] E. Straume, On the geometry and behavior of n-body motions, Int. J. Math. Math. Sci. 28 (12) (2001) 689–732.
[16] E. Strömgren, Connaissance actuelle des orbites dans le problème des trois corps, Bull. Astron. 9 (1935) 87–130.
[17] E. Strömgren, Eine Klasse unsymmetrischer librationsähnlicher periodischer Bahnen im Problème Restreint und ihre Entwicklungsgeschichte (Klasse n), Publ. Copenhagen Obs., vol. 94, 1934.
[18] K.F. Sundman, Recherches sur le problème de trois corps, Acta Soc. Sci. Fennicae 34 (1907) 144–151.
[19] K.F. Sundman, Mémoire sur le problème de trois corps, Acta Math. 36 (1912) 105–179.
[20] A. Wintner, Grundlagen einer Genealogie der periodischen Bahnen im restringierten Dreikörperproblem, Math. Z. 34 (1931) 321–402.